7-30-2013

Improving the Surface-Ground Water Interactions in the Community Land Model: Case Study in the Blue Nile Basin

Di Wu
diwu13@gmail.com

Improving the Surface-Ground Water Interactions in the Community Land Model: Case Study in the Blue Nile Basin

Di D. Wu

Master of Science, University of Connecticut, 2013

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

at the

University of Connecticut

2013

APPROVAL PAGE

Masters of Science in Engineering Thesis

Improving the Surface-Ground Water Interactions in the Community Land Model: Case Study in the Blue Nile Basin

Presented by

Di D. Wu

Major Advisor

Emmanouil N. Anagnostou

Major Advisor

Guiling Wang

Associate Advisor

Mekonnen Gebremichael

University of Connecticut

2013

ACKNOWLEDGEMENTS

First, I would like to thank my major advisors, Dr. Emmanouil N. Anagnostou and Dr. Guiling Wang, for their continuous support, guidance, and patience over these past two years. I am truly grateful for the opportunity to work with these two people. I am also grateful for the assistance and support from my committee member, Dr. Mekonnen Gebremichael.

I would also like to thank Matteo Zampieri and Thymios Serpetzoglou for their CLM3.5 parameterization that was utilized in this study. I also appreciate their help in installing and setting up the model during my first few weeks here at UConn. I also thank Dr. Semu Moges and the Addis Ababa Institute of Technology for their contribution of in-situ precipitation station data.

I want to also thank the entire Civil and Environmental Engineering Graduate Department here at UConn for the support they have given me. Specifically, I would like to thank Rui Mei and Shanshan Sun for their supervision in initial CLM runs, as well as their aid in netCDF conversions. I would also like to thank Dana Parr for additional help and model debugging.

Lastly, I would like to thank my family for all of the support they've given me over these past years: mom, Selena, and Sarah, thank you.

ABSTRACT

Soil moisture is a key water cycle parameter, known to have a positive feedback on precipitation, namely, an increase in soil moisture would increase net radiation and latent heat flux and decrease sensible heat flux and consequentially lower the boundary layer height and increase moist static energy, which eventually leads to precipitation increase. Arguably, land surface models that simulate land surface processes and the surface fluxes to the atmosphere do not capture adequately the spatial variability of soil moisture, particularly over land surface areas with complex topography. A parameterization is applied in this study to the Community Land Model (CLM) 3.5 in an effort to correct for the spatial bias of soil moisture and understand the consequential effects on the simulated water cycle parameters. This parameterization includes a groundwater recharge term from surface water. While CLM contains a river transport model to close the water budget, its current version neglects this groundwater re-infiltration term. Using satellite soil moisture data over the Blue Nile basin, this parameterized term is shown to have a positive correlation to contributing area, which is defined at each grid cell and represents the number of grid cells whose surface drainage accumulates at that local grid cell. With the new parameterization applied to CLM, soil moisture, soil temperature, evapotranspiration flux, water table depth, and vegetation water content all showed significant differences from the control CLM run (without the parameterization) at 95% confidence level. The differences in the spatial distribution of these variables are expected to affect precipitation simulations from regional climate modeling. As Ethiopia is a region that has one of the greatest inter-annual and seasonal precipitation variability globally, the ability to forecast long-term and regional climate predictions is essential. This would allow for optimal reservoir operations including buffering of water resources during times of drought.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SECTION 1: INTRODUCTION

Ethiopia is home to the Blue Nile, located between 34-40°E longitude and 7-13°N latitude. The country houses the largest river network in Africa, in terms of drainage area and mean annual discharge. Despite the fact that the Blue Nile is the main tributary to the Nile River, much of the area's hydrology is not well known (Yilma and Awulachew, 2009). Local authorities are reluctant to publically offer Blue Nile information due to the low international significance of the region, which is fueled by low population density, complex topography, and remote basin location (Conway, 2000). Additionally, Ethiopia is facing one of the greatest hurdles a developing nation will come across, its water resource management. For nations such as Ethiopia, their GDP growth and decay is strongly tied to water availability, as a large portion of their income comes from rain fed agriculture. For Ethiopia, irrigated agriculture only constitutes less than five percent of their total agriculture (World Bank, 2006). Water resource management in this region is challenging due to the large inter-annual and seasonal precipitation variability that is among the strongest globally (Yilma and Awulachew, 2009). With such volatile precipitation, the interactions between climate variability and water cycle dynamics in this region must be better understood. For this reason, climate models should be used to forecast the hydrological variability and, ultimately, be used for long-term and regional climate predictions. This would hopefully allow for early warning and buffering of water resources during times of drought, leading to economic stability and development.

Before regional climate models can be used to simulate the feedback between atmospheric processes and the land surface, they are often initially driven in offline one-way land surface mode for calibration purposes (Zabel et al. 2012). While the regional climate models will often account for groundwater reinfiltration from surface water explicitly at river and watershed scales, the land surface models (LSMs) will usually neglect this term due to lack of

spatial resolution to simulate reinfiltration along river channels (Zampieri et al. 2012). Although many LSMs contain a river transport model (RTM) to simulate fresh water flow from rivers to oceans, thereby closing the water budget, these RTMs also neglect the groundwater reinfiltration term. The importance of modeling this surface water-groundwater interaction has been shown in several studies to influence local precipitation variability (Kingston et al., 2009; Wedgebrow et al., 2002). Additionally, the amount of groundwater storage will greatly affect levels of soil moisture, a key link to the land-atmosphere coupling. The strength and importance of land-atmosphere coupling has been emphasized in many current works (Wang et al. 2007; Wang et al. 2009; Mei and Wang, 2012). These works emphasize the significance of accurately modeling soil moisture, as the positive feedback between soil moisture and precipitation will be essential to capture in a region such as the Blue Nile. By accounting for the reinfiltration of surface water into groundwater, the wetter soils will tend to favor the generation of precipitation by increasing the amount of evapotranspiration. This results in a positive feedback loop where: soil moisture impacts evapotranspiration, evapotranspiration impacts precipitation, and precipitation impacts soil moisture.

Previous attempts at examining the water cycle variables throughout the Blue Nile's extensive river network have proved difficult due to the complex terrain and lack of adequate in-situ data at sub-basin and sub-daily scales (Kim and Kaluarachchi, 2008; Setegn et al. 2009; Sutcliffe and Parks, 2009). Soil moisture values should have large variations over short distance scales due to the complex elevation of the basin, regardless of precipitation amount; areas with higher elevation are expected to have smaller soil moisture values than areas with lower elevation. Since soil moisture is known to influence both local convection and large-scale atmospheric circulations, failure to accurately model the variations in soil moisture will cause

many biases when ultimately coupling the LSM back to a regional climate model (Wang et al. 2007). Even though previous studies within the Blue Nile Basin are able to make short-term predictions, they still lack the ability to predict the hydrologic variability at fine spatial resolution, and at seasonal times-scales or longer.

In Zampieri and coauthors paper in 2012, they proposed a parameterization to be implemented with the Community Land Model 3.5 (CLM) (http://www.cgd.ucar.edu/tss/clm/) that would account for this surface water to groundwater reinfiltration term. This specific LSM was chosen due to the inclusion of rivers and ground water schemes in its RTM, which uses the TOPMODEL approach. As a short summary, the TOPMODEL approach determines per grid cell, the flux of freshwater that will leave that grid cell (Niu et al. 2005). The parameterization introduced in Zampieri's paper focuses on the opposite process which the TOPMODEL approach is based off of, that is, the amount of freshwater flux that will return to each grid cell due to reinfiltration. This surface-groundwater parameterization was introduced to correct the spatial bias of important hydrological parameters, such as soil moisture and evapotranspiration, that CLM 3.5 was seen to underestimate. When the modified CLM was run over Oklahoma (a region that has low-to-moderate orography and river network complexity), the parameterized CLM resulted in statistically significant improvements in many of the water cycle variables – specifically, soil moisture, water table depth, evapotranspiration, and land surface temperatures – at varying depths when comparing model results to station data spread throughout Oklahoma (Zampieri et al. 2012).

In this paper, the method of introducing a groundwater reinfiltration term within the CLM routing model is applied to the Blue Nile river basin to examine the effect of increased model soil moisture spatial variability on hydrological processes in a river basin with complex terrain.

CLM 3.5 will also be used in this study instead of the most recent version. By keeping the same version as the previous study, this will constrain the differences to only complex terrain and extensive river network. Similar to the previous study, different parameterizations of CLM runs will be compared to satellite data at 0.05° resolution (compared to the previous resolution of 0.25°). This finer spatial resolution allows for improved spatial variability analysis of water cycle variables.

This paper is organized as follows: Section 2: STUDY AREA, MODEL DESCRIPTION, AND DATAincludes a description of the study area, the model, and the data used for forcing, validation, and optimization. Section 3: MODEL SETUP describes model setup, including initialization, precipitation bias correction, and parameter validation. Section 4: RESULTS OF MODIFIED CLM includes the results of the modified CLM, specifically focusing on comparisons to satellite data, comparisons to a control initialization of CLM, and importance of results. Finally, Section 5: CONCLUSION contains the conclusion with closing remarks.

SECTION 2: STUDY AREA, MODEL DESCRIPTION, AND DATA

2.1 STUDY AREA

The Blue Nile Basin is located in Ethiopia, between 34-40°E longitude and 7-13°N latitude. It is an area that is characterized by complex mountainous terrain, as seen in Figure 1, with elevations ranging from 329m to 4248m. While the mean elevation of the basin is 1698m, the east and west sides vary greatly in their orography. The east side is much mountainous with a mean elevation of 2100m and standard deviation of 568m. The west side of the basin has much lower elevations but slightly larger variation, a mean of 1290m and a standard deviation of 579m. Southwesterly winds throughout Africa transport moisture from the Congo Basin to the Blue Nile Basin. This moist air will precipitate as it encounters the elevation gradient of the Blue Nile Basin; the increase in elevation from west to east results in larger precipitation amount concentrated over the plains in the west side. Figure 2 shows the cumulative precipitation throughout the region from 2000 to 2010 using National Meteorological Agency of Ethiopia's (NMA) station precipitation data (specifics of this dataset are described in Section 2.3.3 In-Situ Precipitation Data). As described previously, more rainfall is witnessed along the west side due to the elevation gradient and moisture convection. While 11 year cumulative precipitation is shown in the figure, rainfall can reach upwards of 2000mm during a ten-day accumulation during the rainy season (Seleshi and Zanke, 2004).

The Blue Nile Basin has a major rainy season between the months of June through September. These rainy months are a part of the larger east African monsoon season stemming from a shift in the Intertropical Convergence Zone (ITCZ) northward (Conway, 2000). Before the onset of the rainy season, much of the North African land is predominately dry during the months of March through May. This causes a warming of the land and atmosphere within the

region, resulting in a shift of the ITCZ northwards from the equatorial zone. The extent in which the northern part of the ITCZ extends will mainly determine the intensity and duration of the rainy season. However, high pressure systems stemming from the Indian and South Atlantic Oceans as well as thermal lows within Arabia and Sudan will allow for moisture to condense in Ethiopia, further strengthening the precipitation amount during the rainy season (NMSA, 1996; Seleshi and Zanke, 2004). The El Nino (La Nina) Southern Oscillation phenomenon will also result in drier (wetter) than normal conditions. The rainy season persists until the jet stream in Northeast Africa is restored, causing the ITCZ to return south in late September or early October (Conway, 2000). While the northern reach of the ITCZ varies year-to-year, greatly affecting the duration and intensity of the rainy season, the Blue Nile Basin has not experienced significant declines in precipitation amount from 1965 to 2004 (Seleshi and Zanke, 2004). In fact, Seleshi and Zanke found that, unlike other parts of Africa, no significant trend in precipitation accumulation and precipitation duration have been noted over the Blue Nile Basin. This long term stability of rainfall implies that there may also be minimal effects of climate change on precipitation over this area into the early-21st century.

The mountainous terrain of the region results in an extremely complex river network. Figure 3 shows the logarithmic scale of contributing area of the region. In this instance, contributing area is defined at each grid cell, and represents the number of grid cells whose surface drainage is accumulating in the local grid cell. Contributing area is derived from digital elevation data of the region and corresponds with the flow direction field. Furthermore, logarithmic scale is used due to the large range of contributing area values. The Blue Nile River begins at Lake Tana, located in the northern part of the basin shown in Figure 3, and travels southeastwards and finally westwards. The outlet of the Blue Nile Basin, indicated by the red

circle in Figure 3, is located in the northwest of the area and has a logarithmic contributing area value of 3.73; this value of contributing area corresponds to a drainage area of approximately 142,000km^2.

2.2 MODEL

The model used in this study is the Community Land Model (CLM) and is a land surface model for the Community Earth System Model and the Community Atmosphere Model (http://www.cgd.ucar.edu/tss/clm/). The model was developed as a collaborative project between scientists working in the National Center for Atmospheric Research and the university community. CLM is able to be run with a hydrological implementation, which includes rivers and ground water schemes in its river transport model (RTM). The RTM uses the TOPMODEL approach, a one-dimensional representation of a soil column. Water can travel only vertically within the soil column; it is removed by root uptake, evaporation, or runoff (Niu et al. 2005). Runoff occurs at the top three layers of the soil column, as a sum of saturated and unsaturated areas (Oleson et al. 2004). CLM was specifically chosen in the Zampieri et al. (2012) study because of the river and groundwater schemes that facilitates the implementation of different parameterizations. The same model is used here to constrain the differences between this study and the previous one to only complex terrain and extensive river network.

The RTM uses a grid-based, finite difference linear transport scheme to route water from each grid cell to its downstream neighboring grid cell. The total runoff from CLM is accumulated until the RTM is utilized at the end of each time step. At this point, the total runoff is the sum of surface runoff, sub-surface drainage, and liquid runoff from lakes, glaciers, and wetlands (Oleson et al. 2004). The surface runoff component consists of overland flow due to

saturation excess and infiltration excess mechanisms. The linear advection routing scheme in CLM-RTM is governed by the following equation:

$$\frac{dS}{dt} = Q + \sum R_{in} - R_{out} \tag{1}$$

where $\frac{dS}{dt}$ is the temporal variation of water storage, and R and Q are runoff and recharge terms, respectively (Leung et al. 2011).

For the proposed parameterization of surface-groundwater interaction, the interest lies specifically in the groundwater storage equation shown below:

$$\frac{dW_A}{dt} = Q - R_G \tag{2}$$

where W_A is the storage of groundwater in the unconfined aquifer, Q is the recharge of water from unsaturated soil, and R_G is runoff generated from excess water; this equation is the simplified form with the assumptions of constant recharge rate and unfrozen soil (Zampieri et al. 2012). During the parameterization, Equation (2) will be modified to account for the recharge from river water, Q_R:

$$\frac{dW_A}{dt} = Q - R_G + Q_R \tag{3}$$

Since soil moisture is known to be positively correlated with contributing area – approaching a grid cell with larger contributing area (closer to river) will result in higher values of soil moisture – the amount of surface water recharge into groundwater can be parameterized as a function of river water volume over a specific grid cell:

$$Q_R = aW_R^b \tag{4}$$

where W_R is the total river volume in the grid cell as computed by the RTM (Zampieri et al. 2012). The specific values for the parameters a and b will be discussed in Section 3.3 Parameterization Validation. Equation (4) is very general and indicates that the recharge of

groundwater increases as the amount of river water over the cell increases, a notion that is consistent with common sense. This groundwater reinfiltration term is expected to be positively correlated with contributing area. Additionally, this recharge of groundwater will result in increased soil moisture levels.

In the previous study over the Oklahoma region, station and gauge soil moisture values were used to confirm the positive correlation between soil moisture and contributing area. However, lack of adequate station and gauge data within Ethiopia forces the use of satellite data for this confirmation. NASA WindSAT data was chosen from different satellite products for this comparison (details found in Section 2.3.4 Satellite Verification Data). Figure 4 shows both the contour of WindSAT soil moisture values averaged from 2005 to 2010, as well as a scatter plot of the soil moisture values compared to contributing area. The river network is also outlined in the contour plot and, when combined with the contributing area contour shown in Figure **3**, soil moisture is seen to increase from east to west. The west side of the Blue Nile Basin not only has larger contributing area values, but also receives more rainfall from the moisture transported from the Congo Basin. Since the WindSAT product has a spatial resolution of 0.25°, the contributing area values originally at 0.05° were aggregated to that resolution. The scatter plot in Figure 4 takes each soil moisture value and plots it against the corresponding contributing area value. A fitted line is also plotted with a slope of 0.028, confirming a significant positive correlation between soil moisture and contributing area. Furthermore, this slope value is statistically significant at 95% confidence when preforming a one-tailed t-test with a null hypothesis of no slope. Assuming no positive correlation between the two variables, we expect either no slope or insignificant slope. Even at this coarse resolution, the positive correlation

witnessed between soil moisture and contributing area allows for the groundwater recharge term shown in Equation (4) to be used for parameterization purposes.

The varying resolutions at which CLM data will be examined are shown in Figure 5. The fine mesh overlaid represents the 0.05° resolution in which the model is forced and run. Comparisons between the modified CLM and the original CLM will be compared at this resolution. At this 0.05° resolution, there are 114 x 114 grid cells within the region. The larger gray and white squares in Figure 5 represent the 0.25° resolution that the satellite products are available. The CLM results will be aggregated to this coarser resolution whenever model results are compared against satellite data. This coarser resolution results in only 24 x 24 grid cells within the region.

2.3 DATA

2.3.1 ATMOSPHERIC FORCING DATA

CLM 3.5 in the control implementation (CLM-Control), which is without river routing, needs to be forced by atmospheric data and land data. There were no continuous in-situ datasets throughout the complex terrain of the Blue Nile Basin that are able to provide all the atmospheric variables that CLM requires, specifically: specific humidity, precipitation, pressure, radiation flux, temperature, and wind. Thus, the listed variables were taken from the National Centers for Environmental Prediction (NECP) Climate Forecast System Reanalysis (CFSR) dataset. The reanalysis dataset is created through assimilation of a wide range of observational measurements and model results. Observational measurements range from station data, aircraft data, and satellite data which are merged through 3DVAR and 4DVAR techniques with atmosphere, land, and ocean models (Saha et al. 2010). This new reanalysis product was shown to be considerably

more accurate than previous global reanalysis datasets created by the NCEP. Additionally, the dataset is more comprehensive as it includes the newly improve ocean and sea ice models, allowing for higher spatial and temporal resolution (Saha et. al 2010). This results in a complete global CFSR dataset available from January 1979 to March 2011, available at varying resolutions of 0.3°, 0.5°, 1.0°, and 2.5°. Hourly data is provided by either combining the analysis and one-hour through five-hour forecasts, or the one-hour through six-hour forecasts, for each initialization time (Environmental Modeling Center, 2010). The 0.3° resolution data was used for this study, with dates ranging from January 2000 to December 2010. The data was interpolated to 0.05° resolution using inverse distance weighted technique. Only surface height was considered for variables that had many atmospheric levels of data, such as wind, pressure, and humidity. Following the procedure in Zampieri et al. 2012, the first five years of data (2000 to 2004) are used as model spin-up so that the model results no longer depend on initial conditions. Thus, only the last six years from 2005 to 2010 are used in this analysis.

2.3.2 DIGITAL ELEVATION DATA

Mentioned previously, CLM can be run with river routing which requires the additional input of elevation and flow direction field. The HYDRO1k digital elevation data from the USGS Earth Resources Observation and Science Center was used for the region of Ethiopia (https://lta.cr.usgs.gov/HYDRO1K). The data is topographically derived from 30 arc-second elevation data of the world and geo-referenced to 1km resolution. The Blue Nile Basin was extracted from the global elevation dataset and is shown in Figure 1 at the native 1km resolution. The resolution was aggregated to 0.05° resolution and ArcMAPs hydrological toolbox was utilized to determine contributing area and flow directions. These two fields would then also be used as additional forcing for the routing component of CLM.

2.3.3 IN-SITU PRECIPITATION DATA

Due to the complex terrain of the Blue Nile Basin, having accurate precipitation forcing is very crucial in obtaining accurate model results. Unfortunately, very scarce in-situ precipitation data exists for the region that contains fine spatial and temporal resolution. Thus, the model was initially forced with the CFSR precipitation dataset. Since satellite soil moisture values will be used for model evaluation, the precipitation forcing must be unbiased as rainfall has the largest effect on hydrological variables. The numerical model that the CFSR dataset uses to stitch gauge results together also often has significant errors on precipitation determination and provides an additional reason for confirming if there is bias or not. The CFSR reanalysis precipitation was compared to in-situ station data from the National Meteorological Agency of Ethiopia's (NMA) station precipitation data (http://www.ethiomet.gov.et/stations/information) provided by Addis Ababa Institute of Technology. The NMA data is obtained through 43 station measurements of rainfall accumulation during ten-day periods. Majority of these stations are clustered near Lake Tana in the north of the Blue Nile Basin with additional stations in the southeast where there is milder terrain. The station data is interpolated over the entire region resulting in 0.05° spatial resolution, with three data points per month at each grid location. The dataset begins at 1983 and persists to the present. Only the years of 2000 to 2010 were used to match the atmospheric forcing time scale. The process of applying the NMA station data to bias correct the CFSR data is discussed in Section 3.2 Precipitation Bias Correction.

2.3.4 SATELLITE VERIFICATION DATA

CLM soil moisture results will be compared to NASA WindSAT satellite data for validation and calibration purposes. This specific satellite dataset was chosen over Moderate

Resolution Imaging Spectroradiometer (MODIS) and European Space Agency (ESA) due to an extra channel used specifically in the WindSAT product (Li et al. 2010). WindSAT retrievals use polarized 10-, 18-, and 37-GHz channels that are stitched together using maximum likelihood estimate to improve accuracy of the variables provided. Previous studies by NASA had validated the lower frequency channels of 10- and 18-GHz. The new 37-GHz channel aims to offer greater soil moisture sensitivity in varying amounts of vegetation: in areas of low vegetation, the frequency offers high sensitivity; in areas of moderate-to-high vegetation, the frequency offers a linear relationship between vegetation water content brightness and land surface temperature values. One drawback of the 37-GHz channel is that detection is affected by water vapor and clouds in the atmosphere. However, this effect is much less significant over heavily vegetated land such as the Blue Nile Basin. The WindSAT products are obtained from 2000 to 2010 at 0.25° spatial resolution. The satellite offers a daily measurement taken at 00UTC. When comparing CLM results to the WindSAT product, the CLM results are aggregated to the 0.25° resolution, with only one 00UTC result per day from model results.

SECTION 3: MODEL SETUP

3.1 CLM INITIALIZATION

CLM-Control (without river routing) and CLM-Hydro (using the parameterized river transport model) are run in the 0.05° resolution mesh shown in Figure 5. When running CLM-Control, the river routing modules are not utilized. Both of the initializations were started on January 1st, 2000, and lasted 11 years, stopping on December 31st, 2010. Using the procedure in the Zampieri et al. 2012 study, the first five years are used as model spin-up so that model results no longer depend on initial conditions. Thus, analysis and results will only be examined from 2005 to 2010. In addition to the variables mentioned in Section 2.3.1 ATMOSPHERIC FORCING DATA, CLM also requires vegetation, soil type, and texture data. These additional forcing variables were derived from Moderate Resolution Imaging Spectroradiometer (MODIS) products. CLM-Hydro requires additional forcing of digital elevation data and flow direction field as described in Section 2.3.2 DIGITAL ELEVATION DATA.

The water cycle variables that this study will focus upon are: ground surface temperature, near-surface soil moisture, root zone temperature, root zone soil moisture, vegetation water content, evapotranspiration flux, and finally water table depth. The output variables of soil moisture and soil temperature are resolved on a vertical column grid consisting of 10 unevenly spaced layers. Levels closer to the surface have higher resolutions, with the lowest soil layer reaching a depth of 3.43m. The different layers have thicknesses of 1.75, 2.76, 4.55, 7.5, 12.36, 20.38, 33.60, 55.39, 91.33 and 113.7cm (Niu et al. 2007). Specifically, the first three levels are used for ground surface temperature and soil moisture. This corresponds to satellite measurements of soil moisture, as passive microwave retrievals will penetrate the ground approximately 10cm (Li et al. 2010). Levels three through ten are used for root zone soil

moisture and temperature values. A weighted vertical interpolation method is applied over the model soil layers that are within the top 1m of the soil column following the procedure described in Kumar et al. 2009, where the weights are extracted based on the thickness of the soil layer.

3.2 PRECIPITATION BIAS CORRECTION

Initial CLM-Control results were very inconsistent with all satellite products used, whether it was Moderate Resolution Imaging Spectroradiometer (MODIS) data, European Space Agency (ESA) data, or NASA data. Among the forcing variables, precipitation was known to have the greatest effect on water cycle parameters. As mentioned in Section 2.3.3 In-Situ Precipitation Data, the CFSR precipitation data will be compared to in-situ data provided by the National Meteorological Agency of Ethiopia's (NMA). To compare the two precipitation datasets, the CFSR dataset first had to be converted from rain rate to ten-day periods of rainfall accumulation, matching the same temporal resolution as the NMA dataset. The NMA rainfall was available at the same resolution as model forcing, being 0.05°. Afterwards, an initial bias ratio was found using (5)(5 below:

$$Initial\ Bias_{i,j}(t) = \frac{NMA_{i,j}(t)}{CFSR_{i,j}(t)} \tag{5}$$

with a value found for each grid cell and each time measurement. After calculating this initial bias ratio, it became evident that the CFSR precipitation data exhibited some type of bias. Following further inspection, this bias could be separated into three different instances: conditional bias, fraction missed precipitation, and fraction false precipitation.

The conditional bias and coefficient of variation for the wet season months are shown in Figure 6 and Figure 7 respectively. In this case study, conditional bias is calculated with the following equation:

$$Cond.\ Bias_{i,j} = \mu_{i,j}^{log(Bias)}, \quad NMA \geq 1mm\ and\ CFSR \geq 1mm \tag{6}$$

Similarly, the coefficient of variation is calculated by:

$$CV_{i,j} = \frac{\sigma_{i,j}^{log(Bias)}}{\mu_{i,j}^{log(Bias)}}, \quad NMA \geq 1mm\ and\ CFSR \geq 1mm \tag{7}$$

For both these calculations, precipitation values smaller than 1mm per ten-day accumulation were ignored; not only because it is such a low value of rainfall accumulation, but also because the initial bias calculation with such small values would skew the log-normal distribution. For the wet season of the Blue Nile Basin, the CFSR data has minimal bias along the west side of the basin where there are lower elevations. However, the CFSR data has large underestimations of precipitation along the east side of the basin where the terrain is much more jagged with large elevation gradient. While the conditional bias figure shows that the CFSR reanalysis precipitation data has bias that varies with space, the coefficient of variation figure shows that the data also has bias that varies with time. Thus, a single bias ratio cannot be used in attempt to correct the CFSR precipitation; instead, a separate bias ratio for each ten-day accumulation and for each grid cell will be applied.

The fraction of missed precipitation by the CFSR dataset can be seen in Figure 8 and is determined by the following equation:

$$Missed_{i,j} = \frac{\sum_t \left(NMA_{\ i,j}(t) \mid \left[\left(CFSR_{\ i,j}(t) < 1\right)\&\left(NMA_{\ i,j}(t) \geq 1\right)\right]\right)}{\sum_t NMA_{\ i,j}(t)} \tag{8}$$

Specifically, the summation on the numerator only occurs when NMA station data detected precipitation larger than 1mm during ten-day accumulation, whereas the CFSR data does not. This $Missed_{i,j}$ value represents the fraction of NMA station precipitation that is not accounted for by the CFSR dataset. For the wet season, majority of the basin has less than 20% percent of

rainfall missed. Comparable to the conditional bias shown in Figure 6, areas with abrupt elevation changes result in larger fractions of missed precipitation. Nonetheless, during the wet season of the Blue Nile, rainfall can reach upwards of 2000mm in ten-days. This emphasizes the significance of missing greater than 50% of the rainfall as it will result in large changes of hydrological variables during model simulations.

Lastly, the fraction of false precipitation accumulation by the CFSR dataset can be seen in Figure 9 and is determined by the following equation:

$$False_{i,j} = \frac{\sum_t \left(CFSR_{i,j}(t) \mid \left[\left(CFSR_{i,j}(t) \geq 1\right) \& \left(NMA_{i,j}(t) < 1\right)\right]\right)}{\sum_t NMA_{i,j}(t)} \tag{9}$$

where the summation on the numerator only occurs when NMA station data does not detect any precipitation larger than 1mm during ten-day accumulation, whereas the CFSR data reports larger than 1mm precipitation values. Although the fraction of false precipitation is very low and will not contribute much to the wet season bias correction, it should be more significant when applied to the year-round bias correction.

To maintain the climatology of the CFSR precipitation data while retaining the fine-resolution, spatial distribution of precipitation from NMA station data, the final bias ratio will be applied in the following way:

$$Final\ Bias_{i,j}(t) = \begin{cases} 0, & NMA_{i,j}(t) < 1mm \\ \dfrac{NMA_{i,j}(t)}{10}, & \left(NMA_{i,j}(t) \geq 1mm\right)\ and\ \left(CFSR_{i,j}(t) < 1mm\right) \\ \dfrac{NMA_{i,j}(t)}{CFSR_{i,j}(t)}, & other\ cases \end{cases} \tag{10}$$

- Case: $Final\ Bias_{i,j}(t) = 0$. Occurs when NMA station measurement detects less than 1mm of rain in a ten-day accumulated rainfall. $Final\ Bias_{i,j}(t)$ will be applied multiplicatively to

each CFSR data. This accounts for the cases of *false precipitation* detection by CFSR reanalysis data.

- Case: $Final\ Bias_{i,j}(t) = NMA_{i,j}(t)/10$. Occurs when CFSR data displays less than 1mm of rain in a ten-day accumulated rainfall while NMA station data measures anything larger than 1mm. The NMA station measurement will be taken as the actual rainfall. The rainfall accumulation will be converted back into a precipitation rate and will be distributed evenly throughout one-third of the days in a specific month (on average, this will be 10 days). During bias correction, the CFSR precipitation rate will be set directly equal to $Final\ Bias_{i,j}(t)$. This will force the CFSR reanalysis data to equal the NMA station data, rather than applying the bias ratio multiplicatively. The even distribution of precipitation during one of these periods will not affect the hydrological variables significantly since the evaluation of model results is computed on seasonal time-scales. Rather, not accounting for this missed precipitation amount by CFSR would have a larger negative impact on water cycle variables than this even distribution of station precipitation. This accounts for the cases of *missed precipitation* by CFSR reanalysis data.
- Case: $Final\ Bias_{i,j}(t) = NMA_{i,j}(t)/CFSR_{i,j}(t)$. The same format as the $Initial\ Bias_{i,j}(t)$ equation found in Equation (5). This accounts for all other cases of difference between the two precipitation datasets.

This precipitation bias correction ultimately resulted in large changes in CFSR precipitation values that were distributed over the region. While satellite precipitation data could have been used to force CLM, it ultimately would not have made a difference since the same bias correction procedure would have been applied. Whatever precipitation dataset was used for model forcing, it would closely match the NMA station data after the bias correction method.

3.3 PARAMETERIZATION VALIDATION

After correcting the bias of the CFSR precipitation dataset, CLM results began to look comparable to the NASA WindSAT product. The next step would be to optimize the parameterization included to the groundwater recharge term, with equation shown again below:

$$Q_R = aW_R^b \tag{4}$$

Since a positive correlation between soil moisture and contributing area was confirmed in Section 2.2 MODEL, a positive coefficient will be used for parameters a and b. The focus will also be on soil moisture for parameterization validation, as it relates directly to the reinfiltration term assumption. Similar to Zampieri et al.'s (2012) study, two approaches for the values of parameter b will be examined. These two values will be $b = 1$ and $b = 2/3$ and correspond to a recharge term depending on *volume* or *surface* of river water over a grid cell, respectively. For the dimensions within Equation (4) to be equal, parameter a will have units of water infiltration speed when $b = 1$, and units of conductivity when $b = 2/3$.

These parameterizations are intended to remove the bias that many land surface models will have as you approach larger contributing area values. Figure 10 graphs preliminary CLM-Control results of near-surface soil moisture against those from WindSAT data, with averages from 2005 to 2010. These near-surface soil moisture results are obtained from the first three layers of soil levels, and aggregated to 0.25° resolution to match that of the WindSAT product. Only the wet season months, June through September, are used during comparisons to avoid the masking effect frozen soils have on soil moisture satellite results (Wang et al. 2007). Figure 10 was created in a similar format as the previously described Figure 4 scatter plot. Specifically, the ratio was taken between WindSAT and CLM-Control near-surface soil moisture value, at each i

and j location. Afterwards, this ratio is plotted against the corresponding contributing area value at i and j. This results in the ratio-contributing-area plot shown in Figure 10. As seen in the figure, bias increases in CLM-Control results with increasing contributing area values. Assuming a perfect match between model results and satellite data, there should be no increase or decrease in ratios with varying contributing area. CLM-Control is shown to underestimate soil moisture at larger contributing areas, which results in greater ratio values. This means that the model exhibits a dry, and therefore warm, bias with respect to the satellite product. A way to numerically obtain a measure of this bias would be to take the slopes for each year and examine their values. An optimal set of parameters values for a and b in Equation (4) should result in little to no slopes.

The values for parameters a and b were evaluated subjectively, with model derived climatology compared against WindSAT satellite products after each parameter set. A wide range of values for parameter a were tested given each case of parameter b. Not all cases of the two parameters are discussed, just the outer ranges of parameter a that show significant changes in model derived climatology. Figure 11 shows four different parameterization schemes in the ratio-contributing-area plot. The different lines represent the average ratio from 2005 to 2010, once again only including wet season months. A line has been drawn across a value of $y = 1$, which is what a perfect match between model results and satellite data should result in. Table 1 tabulates the different sets of parameter values for a and b, as well as the slope average from 2005 to 2010. As seen in both this table and Figure 11, when parameter $b = 1$ the ratios are closer to a slope of zero than when parameter $b = 2/3$. It is interesting to note that any parameterization results in an improvement of soil moisture values. CLM-Control reaches an average ratio of approximately 1.45 at contributing area value of 2.6 as shown in Figure 10. However, the two parameterizations of $b = 2/3$, which perform worse than when parameter

$b = 1$, reach peak ratio values of less than 1.25 as shown in Figure 11. Additionally, changes to parameter a when $b = 2/3$ seem to have little to no effect on soil moisture spatial variability within the region. The values between soil moisture and contributing area for these two parameterizations are nearly overlaid on top of one another, resulting in very similar slopes and resulting plot. From the figure and slopes, the parameterization that produces best results compared to WindSAT products is when $a = 1 \times 10^{-7}\text{s}^{-1}$ and $b = 1$.

Since parameter a has units of inverse time scale in this instance, one way to interpret this value is the time scale necessary for a river-scale induced groundwater anomaly to influence the soil moisture in an area of 0.05° by 0.05° (Zampieri et al. 2012). This value corresponds roughly to a time of 3 months, indicating a process on the same time scale as the wet season duration. While this is not a precise estimate due to the highly parameterized model with specific conditions, it does imply that the timescale corresponds to a process which occurs on the order of several months. Additionally, this parameterization makes sense intuitively when considering the result from the previous Oklahoma study. In the previous study, it was found that the parameterization $a = 1 \times 10^{-9}\text{s}^{-1}$ and $b = 1$ was the slightly better parameterization. Since parameter a corresponds to a time scale, the Blue Nile with its much faster flow and larger river network should have a faster time.

SECTION 4: RESULTS OF MODIFIED CLM

4.1 MODEL RESULTS COMPARED TO WINDSAT DATA

The analysis conducted on the CLM results will only focus on the months of June through September when rainfall is the most significant and the masking effect of frozen soil on soil moisture infiltration is negligible (Wang et al. 2007). Figure 10 shows the increasing underestimation of CLM-Control soil moisture as contributing area increases based on the comparison of model results with the WindSAT data. Model soil moisture values are aggregated to 0.25° resolution to match that of the satellite product. While there is variability between the six years shown, the overall trend is that there is an almost exponential increase in soil moisture underestimation at large contributing areas. This is consistent with the general notion that many land surface models do not account for the reinfiltration of surface water to groundwater, resulting in the inability to accurately capture the spatial distribution of soil moisture.

In the previous section, the optimal parameterization for the region was determined with values of $a = 1 \times 10^{-7}\text{s}^{-1}$ and $b = 1$, to be applied to Equation (4); this simulation will further be referred to as CLM-Hydro. Figure 12 graphs the percent improvement of CLM-Hydro near-surface soil moisture when compared to CLM-Control. This figure was created by first generating a graph similar to Figure 10, but with WindSAT / CLM-hydro ratio on the y-axis. Afterwards, the WindSAT / CLM-Hydro ratios were subtracted from the WindSAT / CLM-Control ratios. The result reflects the relative improvement of soil moisture simluation from one scheme to the next. The dark black line represents the average of the six years, with thinner colored lines representing the improvement in individual years. Soil moisture is once again the focus of this comparison due to its positive correlation with the reinfiltration parameterization term. It can be seen that, while fluctuating throughout the years and contributing areas, larger

contributing areas will result in a larger improvement of soil moisture. This improvement reaches a peak of just over 10% at the largest contributing area values. While the satellite data is only available at 0.25° resolution, a much smoother percent improvement graph is expected if the satellite data was available at a finer resolution. Much of the model variability of soil moisture was lost as results were aggregated from the 0.05° resolution to the 0.25° resolution.

4.2 MODIFIED CLM COMPARED TO CONTROL CLM

After confirming improvement of the CLM-Hydro parameterization against the WindSAT soil moisture product, comparisons can now be made between CLM-Hydro and CLM-Control. These assessments are now made at the model resolution of 0.05°. This higher resolution comparison should allow the new parameterization to capture the spatial variability of the water cycle parameters. Figure 13 through Figure 17 show contours of CLM-Hydro / CLM-Control ratios of root zone soil moisture, near-surface soil moisture, water table depth, evapotranspiration flux, and vegetation water content, respectively. Figure 18 and Figure 19 show contours of the difference between CLM-Hydro and CLM-Control for root zone soil temperature and near-surface temperature. All contours are the mean values from 2005 to 2010. As mentioned in Section 3.1 CLM INITIALIZATION, near-surface temperature and soil moisture values use the first three levels from the CLM soil column. Levels three through ten are used for root zone soil moisture and temperature values. In either case, a weighted vertical interpolation method is used over the model soil layers that are within the top 1m of the soil column following the procedure described in Kumar et al. (2009).

The ratio contour plots have their increments split into nine, unevenly spaced intervals around a value of one. The uneven increments were chosen due to the log-normal distribution

that occurs when ratios are taken. Besides the increments shown, any ratio value less than 0.70 is grouped into the lower interval, and any ratio value greater than 1.40 is grouped into the upper interval. Any ratio between 0.95 and 1.05 are not considered to be a statistically significant improvement of variables and thus, are graphed in white. For all ratio contour plots, colors approaching red occur when CLM-Hydro produces larger soil moisture values than CLM-Control. In the specific case of water table depth, the ratio is instead taken as CLM-Control / CLM-Hydro; the reason for this will be discussed later on in this section. Likewise, any colors approaching blue occur when CLM-Hydro produces smaller soil moisture values than CLM-Control.

Figure 20 combines all the water cycle variables into one line graph. This figure was created in a similar format as the previously described Figure 4 scatter plot. Specifically, the ratio (or difference) was taken between CLM-Hydro and CLM-Control for each separate variable at each *i* and *j* location. Afterwards, this ratio is plotted against the corresponding contributing area value at that particular *i* and *j*. Due to the temperature scale discrepancy mentioned in the previous paragraph, Figure 20 contains two separate y-axes: the left-side is the *ratio* between CLM-Hydro and CLM-Control, and the right-side is the *difference* between CLM-Hydro and CLM-Control. The two temperature lines – orange representing ground surface temperature and cyan representing root zone soil temperature – are graphed on the second y-axis as a difference. All other variables are all graphed on the first y-axis of ratios. Furthermore, the ratio value of *one* and difference value of *zero* has been centered to show a line of "no improvement". The x-axis of logarithmic contributing area values is also stopped at 2.4, pertaining to a drainage area of approximately 6,275km^2. Larger values are excluded do to the rapid increase in soil moisture as you approach rivers and lakes, which will result in extremely large (or small) ratios of variables

that will skew results. Slopes can be calculated from the line graphs between contributing area values of 0.7 to 2.4 that would emphasize the relative change in water cycle variables from CLM-Hydro to CLM-Control. This starting point of 0.7 was chosen by examining where the water cycle parameters begin exhibiting an opposite trend; as seen in Figure 20, this occurs roughly between contributing area values of 0.6 to 0.8. This opposing trend occurs due to the additional elevation and flow direction field required as input for CLM-Hydro. By accounting for elevations and flow accumulation, areas with higher elevations will now have lower modeled soil moisture than in CLM-Control. This resultant decrease in soil moisture will correlate to the other changes in water cycle variables for these locations with higher elevations. Similarly, the opposite effect will occur in areas where puddling exists resulting in increased soil moisture levels between CLM-Hydro and CLM-Control. Both of these cases are essential for better capturing the spatial variability of water cycle parameters within the region. A table of the slopes can be found in Table 2, with the slope of each year separately shown, as well as the mean and coefficient of variation for all six years. These slope values give the relative improvement of each variable and are unit-less. Larger absolute value of slopes corresponds to a greater change from CLM-Hydro to CLM-Control. Coefficient of variation gives insight to the degree of variation each separate year has compared to the mean.

Root zone soil moisture and near-surface soil moisture contours are shown in Figure 13 and Figure 14 respectively. For both of these contours, the entire river network can be seen in red with corresponding ratio values greater than or equal to 1.40. For the near-surface soil moisture contour, majority of the basin has increases in soil moisture. This is a direct result of the reinfiltration of groundwater due to surface water parameterization that has been included in the CLM-Hydro implementation. The increase in both soil moisture variables reaches a maximum

when approaching the main river location. The regions in which soil moisture is lower in CLM-Hydro than in CLM-Control are shown in shades of green and blue. The distribution may seem random on the contour maps, but is caused by the inclusion of elevation and drainage paths when running CLM-Hydro. This effect is more easily seen in Figure 20, where the decrease in soil moisture begins at extremely low contributing area values, and reaches a maximum at a value of 0.7.

As seen in Figure 20, water table depth ratio decreases to a value very close to zero when approaching a contributing area value of 2.4. This is due to the way CLM defines water table depth. The value for the variable decreases as the water table becomes closer to the surface as a result of increased soil moisture. As a consequence, grid cells close to the main river network have water table depth values of nearly zero. Thus, taking the ratio between CLM-Hydro / CLM-Control also results in values close to zero. However, when the water table depth contour is shown in Figure 15, the opposite ratio is taken: CLM-Control / CLM-Hydro. This inverse ratio is taken for visual consistency. On all ratio contour plots, red coloring indicates that CLM-Hydro is producing larger values of soil moisture. As stated above, due to the definition of water table used by CLM, increasing values of soil moisture result in decreasing values of water table depth. Only the water table depth contour utilizes the inverse ratio, or the line graph in Figure 20 would approach positive infinity if this same inverse ratio was used. Examining the contour of water table depth, it gets shallower throughout majority of the Blue Nile Basin besides the extremely mountainous east side of the region. Additionally, majority of areas with shallower water table depth experience greater than 1.40 ratio increases from CLM-Control values.

Compared to the changes in soil moisture and water table depth, the change in evapotranspiration flux from CLM-Hydro to CLM-Control is much weaker. In Figure 20, the

change of total evapotranspiration is approximately half the strength of the near-surface soil moisture change. Even the contour shown in Figure 16 displays an almost even amount of areas where CLM-Hydro results in larger evapotranspiration, to areas where CLM-Hydro results in smaller evapotranspiration. One possible reason for this minor change in evapotranspiration can be attributed to the difference in limiting factor between the two CLM implementations. Figure 21 takes the evapotranspiration flux graph from Figure 20, and separates the total flux into canopy evaporation, canopy transpiration, and ground evaporation components. This figure follows the same format as Figure 20 and contains the average of ratios from 2005 to 2010. As shown in the figure, the major contributor to the increase in total evapotranspiration flux lies with ground evaporation and the other two canopy fluxes are both smaller in CLM-Hydro than CLM-Control. Since the overall magnitude change of total evaporation is not comparable to the changes shown with soil moisture, some limiting factor must be causing the dampening of the signal to propagate to evapotranspiration. When CLM-Control is run, soil moisture limits the amount of evapotranspiration flux over most of the domain. However, the parameterization implemented to CLM-Hydro allows for much more wetter soils in areas of large drainage areas and energy becomes the limiting factor for evapotranspiration flux. The thought behind this justification can be found by how CLM treats vegetated areas such as the Blue Nile Basin, where the total water vapor flux is balanced by vegetative flux and ground flux (Oleson et al. 2004). The ground flux term has a lower resistance term than vegetative flux, with includes both stomatal resistance and air resistance. This explains why ground evaporation flux experiences a large increase almost on the same magnitude as soil moisture increase during the CLM-Hydro implementation. With this energy limitation, an increase in ground evaporation will result in decreases of overall decreases of vegetative flux, which includes both canopy evaporation and

transpiration (Oleson et al. 2004). This energy limitation is also able to describe why the total evapotranspiration contour experiences very minor increases and decrease throughout the basin, when compared to any of the other water cycle parameters.

While vegetation water content increases throughout the basin, the contour shown in Figure 17 displays that this variable does not experience as large of a change as soil moisture. When examining the dependence of vegetation water content ratios and contributing area in Figure 20, a net increase is seen throughout contributing areas. However, this change in vegetation water content is very minimal even as contributing area values range from low to high. This is partially due to the large plain in the southwest that experiences large values of vegetation water content increases. The southwest area actually has lower contributing area values and, as a consequence, will reduce the slope of the curve. Due to the shared vegetation data between CLM-Hydro and CLM-Control, it seems the vegetation water content is more dependent on precipitation distribution than differences in soil moisture. Nonetheless, the parameterization applied to CLM-Hydro did result in overall increased vegetation water content.

Figure 18 and Figure 19 show the contours of near-surface ground temperature and root zone soil temperature, respectively. Mentioned previously, both the contours and the graph of temperatures in Figure 20 are of differences between CLM-Hydro and CLM-Control, instead of the ratios used for the other water cycle parameters. The two contour figures look very similar, both in terms of magnitude and spatial distribution. While root zone soil moisture has lower temperature values compared to near-surface soil moisture by roughly 1°K, both temperature variables show majority of areas that are cooler. The exception to this rule is the plain in the southwest side of the area, where temperatures have instead increased by 1°K. The increase in temperature within this area is due to the evaporative cooling effects of evapotranspiration. Since

the southwest area actually exhibits a decrease in evapotranspiration flux in the parameterized CLM-Hydro, it will result in warmer ground temperatures for that region. These effects can be seen in Figure 20 where temperatures around a contributing area value of 0.7 are warmer, and at higher contributing area values the root zone soil temperature is lower than surface temperature by approximately 1°K. Temperatures are pretty consistent between CLM-Hydro and CLM-Control along the western side of the basin, where there is flatter terrain and more uniform precipitation pattern.

4.3 Importance of Results and Model Coupling

To examine the relative improvements of the spatial variability for each of the water cycle parameters, the slope can be taken between contributing area values of 0.7 to 2.4 from Figure 20. The slope values for each year are tabulated in Table 2, along with the mean and coefficient of variation. By examining both the table and the figure, the different variables can be seen to vary in the amount the parameterization has altered them from CLM-Control. Additional insight can be gained by examining the inter-dependence of water cycle parameters within CLM due to these different degrees of variation. As soil moisture increases from low-to-high values of contributing area, evapotranspiration flux, and vegetation water content all follow a positive correlation. Water table depth and soil temperatures exhibit a negative correlation, decreasing as contributing area increases. Looking at the mean slope values of each water cycle parameter located in Table 2, root zone soil moisture has the largest change when applying this new parameterization. By increasing groundwater storage through reinfiltration by Equation (4), it directly influences root zone soil moisture to have the greatest change with respect to increasing contributing area, resulting in a slope value of 0.2485. Near-surface soil moisture also shows

dramatic improvements with the new parameterization, with a slope of 0.2064. However, due to way CLM treats evapotranspiration, where the top-most-layers of the soil column experience much larger changes compared to the lower layers, the slope of evapotranspiration flux is slightly less than the root zone soil moisture value. It follows intuitively that root zone soil moisture is much more resilient to short term change than near-surface soil moisture. Water table depth also decreases dramatically throughout this new parameterized CLM-Hydro. Water table depth is shown to have a strong negative correlation with changes to soil moisture yielding a slope value of -0.2433. Evapotranspiration flux and vegetation water content both increase with approximately same relative slope, with values of 0.0996 and 0.0839 respectively. Both of these variables also exhibit a positive correlation to changes in soil moisture; however, the magnitude that these variables are affected is much less than the magnitude of soil moisture changes. In the case of evapotranspiration flux, the limiting factor in CLM-Hydro is now energy in the form of water vapor flux. For vegetation water content, the value of change seems to depend more on precipitation patterns or forcing data than the increase in soil moisture. Finally, the relative changes of soil temperature and near-surface temperature are -0.0038 and -0.0021 respectively. While these values indicate weak negative correlations the Kelvin scale of temperature measurement masks the significance of the parameterization, as a soil temperature difference of 2°K is quite noteworthy (Zampieri et al. 2012). Regardless of the magnitude of the slopes, all the variables have statistically significant change at 95% confidence. These tests were all made using a two-tailed t-test with a null hypothesis of no slope. If the parameterization of CLM-Hydro instead resulted in no change from CLM-Control, the calculated slope values should have been close to zero or statistically insignificant.

The parameterization imposed in this study improves the spatial variability of all water cycle variables examined. The analysis conducted on the two different implementations of CLM only accounts for the wet season months of June through September. When investigating the soil moisture-precipitation coupling strength, there are very few areas in the world with modest-to-strong coupling during the dry season (Wang et al. 2007). The previous Oklahoma study also found that the parameterization applied to CLM-Hydro performs better during wet and wet-preceded climate conditions (Zampieri et al. 2012). These serve as additional reasons for analysis performed only on the summer months within the Blue Nile Basin.

The importance of these changes to water cycle parameters is influenced by the direct coupling between soil moisture and future precipitation. Since Ethiopia has one of the greatest inter-annual and seasonal precipitation variability globally, future prediction of precipitation is essential for the country's growth and development (Yilma and Awulachew, 2009). One of the links connecting soil moisture and precipitation can actually be seen from model results: increased soil moisture has resulted in increased evapotranspiration. This evapotranspiration will ultimately influence precipitation through its impact on local convection and large-scale atmospheric circulation. Additionally, the stronger signal shown by root zone soil moisture compared to near-surface soil moisture emphasizes the importance of capturing the spatial variability of the variable, as the residence time of water within the deeper soil is much longer than the residence time of water within the atmosphere.

In a previous study by Wang et al. (2007), the soil moisture influence on future precipitation was examined through its impact on future evapotranspiration. Through a point-to-point correlation between soil moisture and evapotranspiration, it was found that soil moisture did have a significant impact on evapotranspiration in the Blue Nile region. Additionally, this

same area showed moderate-to-strong coupling in Koster's previous research (2004, 2006). These studies accentuate the importance of capturing the spatial variability of soil moisture, especially in regions with complex terrain and large river network, like the Blue Nile Basin.

SECTION 5: CONCLUSIONS

In this study, a parameterization was applied to CLM 3.5 which included an additional term in the groundwater storage equation which accounted for the reinfiltration of surface water. By including this reinfiltration term that is positively correlated with contributing area, it hopes to correct the spatial bias that is present in many land surface models when determining water cycle variables. This parameterization was previously applied in Zampieri et al. 2012 study over Oklahoma, with promising results during the wet season months during the summer, specifically the months of June through August. This new study applies this groundwater-surface water parameterization in the study area of Blue Nile Basin, located in Ethiopia. By utilizing the same version of CLM, the differences between this study and the previous only lie with the complex terrain and large river network that is associated with the Blue Nile Basin.

The groundwater reinfiltration term implemented in CLM-Hydro was first justified by showing the positive correlation between soil moisture and contributing area. This justification was confirmed by utilizing NASA WindSAT soil moisture data at a coarse resolution of 0.25°. By affirming this positive correlation, an optimal parameterization was determined by examining different subsets of the parameterization compared to the same satellite data. After the optimal parameterization was found for the study region, CLM-Hydro results would be compared to CLM-Control, where river routing is not included. By comparing these two implementations at a high resolution of 0.05°, it would serve to confirm the objective that land surface models do exhibit spatial bias of key water cycle variables. Besides the increase in soil moisture expected from the reinfiltration term, the parameterized CLM-Hydro subsequently decreased soil temperatures. The higher soil moisture values resulted in greater soil evaporation, causing greater evaporative cooling and lower surface temperatures. The improvement of root zone soil moisture, which constituted the bottom-most seven layers of the soil column, was larger than the

improvement of near-surface soil moisture, which constituted the top-most three layers. Additionally, this increase in soil moisture raised the water table depth closer to the surface. This higher water table also allowed for increased vegetation water content throughout the area.

While improvements to the spatial variability of the water cycle variables are all statistically significant at 95% confidence, the initial verification and calibration of the parameterization were all based off of the NASA WindSAT soil moisture product. Other satellite products such as ESA and MODIS were viewed, but they either did not match the soil moisture spatial distribution of the region well, or they did not have complete coverage of the region. The WindSAT product was also only available at 0.25° resolution whereas CLM was run at 0.05° resolution. Due to the scarcity of soil moisture data within the region, either having finer resolution satellite products or having more variables to optimize would result in more accurate modeled water cycle variables. Other limitations of the study include the linear routing scheme implemented by CLM-RTM. While this linear routing allows for easy parameterization implementations, it is not as accurate as other, more complex, routing schemes. This linear routing scheme also does not account for water management within the region, although most of this management will be applied to agriculture along the west side of the Blue Nile Basin. As Ethiopia develops further as a country, station and gauge data could ultimately be used to force models instead of having to rely solely on NCAR CFSR datasets. Finally, this parameterization of groundwater reinfiltration could be updated to be included in the newest version of CLM available. These aspects can be explored in future case studies. As a closing statement, the inclusion of the reinfiltration of river water into the ground water storage improved the comparison with satellite data; therefore, this parameterization may be utilized as a method of integrating satellite data into land surface models to provide more realistic results.

REFERENCES

Ahmed, K.A., Wang, G., Silander, J., Wilson, A.M., Allen, J.M., Horton, R., Anyah, R., 2013. "Statistical downscaling and bias correction of climate model outputs for climate change impact assessment in the U.S. northeast." *Global and Planetary Change*, Vol. 100, pp. 320-332

Block, P., and Goddard, L., 2011. "Statistical and Dynamical Climate Predictions to Guide Water Resources in Ethiopia." Journal of Water Resources Planning and Management, Vol. 138, Issue 3, pp. 287-298.

Blöschl, G., and Zehe, E., 2005. "On hydrological predictability." *Hydrological Processes*, Vol. 19, Issue 19, pp. 3923-3929.

Conway, D., 2000. "The Climate and Hydrology of the Upper Blue Nile River." *The Geographical Journal*, Vol. 166, No. 1, pp. 49-62

Das, N.N., Entekhabi, D., and Njoku, E.G., 2010. "An algorithm for merging SMAP radiometer and radar data for high resolution soil moisture retrieval." *IEEE Trans. Geosci. Remote Sens.* Vol. 49, Issue 5, 1504-1512

Environmental Modeling Center, National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, 2010: NCEP Climate Forecast System Reanalysis (CFSR) Selected Hourly Time-Series Products, January 1979 to December 2010. *Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, CO.* [Available online at http://rda.ucar.edu/datasets/ds093.1.]

Koster, R. D., Dirmeyer, P. A., Hahmann, A. N., Ijpelaar, R., Tyahla, L. Cox, P., and Suarez, M. J., 2002. "Comparing the degree of land–atmosphere interaction in four atmospheric general circulation models." J. Hydrometeor, Vol. 3, pp. 363-375.

Koster, R. D., Dirmeyer, P. A., and Guo, Z. C., 2004. "Regions of strong coupling between soil moisture and precipitation." Science, Vol. 305, pp. 1138–1140.

Koster, R. D., and Coauthors, 2006. "GLACE: The Global Land-Atmosphere Coupling Experiment. Part I: Overview." J. Hydrometeor., Vol. 7, pp. 590–610.

Kim, U. and Kaluarachchi, J., 2008. "Application of parameter estimation and regionalization methodologies to ungauged basins of the Upper Blue Nile River Basin, Ethiopia." *Journal of Hydrology*, Vol. 362, Issue 1-2, pp. 39-56

Kingston, D. G., Hannah, D. M., Lawler, D. M., and McGregor, G. R., 2009. "Climate-river flow relationships across montane and lowland environments in northern Europe." *Hydrol. Process*. Vol. 23, Issue 7, pp. 985-996

Kumar, S.V., Reichle, R.H., Koster, R.D., Crow, W.T., Peters-Lidard, C.D., 2009. "Role of subsurface physics in the assimilation of surface soil moisture observations." *J. Hydrometeorol*., Vol. 10, pp. 1534-1547.

Lawrence, P.J., Chase, T.N., 2007. "Representing a new MODIS consistent land surface in the community land model (CLM 3.0)." *J. Geophys*. Res. 112, G01023. doi:10.1029/2006JG000168

Leung, L. R., et al., 2011. "Development of River Routing and Groundwater Models in CLM." *CESM Land Model Working Group Session*

Li, L., Gaiser, P. W., Gao, B.-C., Bevilacqua, R. M., Jackson, T. J., Njoku, E. G., Rudiger, C., Calvet J.-C., and Bindlish, R., 2010. "WindSat Global Soil Moisture Retrieval and Validation." *Geoscience and Remote Sensing*, Vol. 48, Issue 5, pp. 2224-2241.

Luseno, W. K., McPeak, J. G., Barrett, C. B., Little, P. D. & Gebru, G., 2003. "Assessing the Value of Climate Forecast Information for Pastoralists: Evidence from Southern Ethiopia and Northern Kenya." *World Development*, Vol. 31, pp. 1477-1494

Ministry of Water and Energy, 2011. "Tana and Beles Integrated Water Resources Development Project-Design, Development & Support in Implementation & Operation of a Basin Information System." *Final Inception Report*

Niu, G.-Y., Yang, Z.-L., Dickinson, R.E., Gulden, L.E., 2005. "A simple TOPMODEL-based runoffparameterization (SIMTOP) for use in global climate models." *J. Geophys*. Res. 110, D21106, doi: 10.1029/2005JD006111

Niu, G.-Y., Yang, Z.-L., Dickinson, R.E., Gulden, L.E., Su, H., 2007. "Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data." *J. Geophys*. Res. 112, D07103. doi:10.1029/ 2006JD007522

NMSA 1996. "Climate and agroclimate resources of Ethiopia." NMSA Meterological Research Report Series., Vol. 1, No. 1, Addis Ababa.

Oleson, Keith W., et al. 2004. "Technical description of version 3.0 of the Community Land Model (CLM)."

Saha, S., and Coauthors. 2010. "The NCEP Climate Forecast System Reanalysis." *Bull. Amer. Meteor. Soc.*, Vol. 91, pp. 1015-1057.

Seleshi, Y., and Zanke, U., 2004. "Recent changes in rainfall and rainy days in Ethiopia." International Journal of Climatology., Vol. 24, pp. 973-983.

Seneviratne, S.I., Corti, T., Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowsky, B., and Teuling, A.J. 2010. "Investigating soil moisture-climate interactions in a changing climate: A review." *Earth-Science Reviews*, Vol. 99, pp. 125-161

Setegn, S. G., Srinivasan, R., and Dargahi, B., 2009. "Hydrological Modeling in the Lake Tana Basin, Ethiopia Using SWAT Model." *Open Hydrol. J.*, Vol. 2, pp. 49-62

Sutcliffe, J. V., and Parks. 2009. "Hydrological modeling of the Sudd region of the Upper Nile." *hydrol. Sci. Bull.*, Vol. 19, Issue 2, pp. 237 -255

Wang, D., Wang, G., Anagnostou, E.N., 2009. "Impact of sub-grid varability of precipitation and canopy water storage on hydrological processes in coupled land-atmosphere model." *Clm. Dyn.*, Vol. 32, pp. 649-662

Wang, G., Kim, Y., and Wang, D., 2007. "Quantifying the Strength of Soil Moisture–Precipitation Coupling and Its Sensitivity to Changes in Surface Water Budget." J. Hydrometeor, Vol. 8, pp. 551-570.

Wedgbrow, C. S., Wilby, R. L., Fox, H. R., and O'Hare , G., 2002. "Prospects for seasonal forecasting of summer drought and low river flow anomalies in England and Wales." *Int. J. Climatol.*, Vol. 22, Issue 2, pp. 219-236

World Bank. 2006. "Ethiopia: Managing water resources to maximize sustainable growth." *A World Bank Water Resources Assistance Strategy for Ethiopia*

Wood, A.W., Maurer, E.P., Kumar, A., Lettenmaier, D.P., 2002. "Long-range experimental hydrologic forecasting for the eastern United States." *Journal of Geophysical Research*, Vol. 107, Issue D20, pp. 4429

Yilma, A. D. and Awulachew, S. B., 2009. "Characterization and Atlas of the Blue Nile Basin and its Sub basins." *International Water Management Institute*

Zabel, F., Mauser, W., Marke, T., Pfeiffer, A., Zängl, G., and Wastl, C. 2012. "Inter-comparison of two land-surface models applied at different scales and their feedbacks while coupled with a regional climate model." *Hydrol. Earth Syst. Sci.*, Vol. 16, pp. 1017-1031, doi:10.5194/hess-16-1017-2012

Zampieri, M., Serpetzoglou, E., Anagnostou, E.N., Nikolopoulos, E.I., Papadopoulos, A., 2012. "Improving the representation of river-groundwater interactions in land surface modeling at the regional scale: Observational evidence and parameterization apllied in the Community Land Model." *J. Hydro.*, Vol 420-421, pp. 72-84

TABLES

Table 1: Values of slope for each CLM-Hydro parameterization from Figure 11. Each slope value represents the average from 2005 to 2010, only considering months of June through September.

Parameterization Parameters		Slope
$a = 1 \times 10^{-3}$	$b = 2/3$	0.04937
$a = 1 \times 10^{-5}$	$b = 2/3$	0.04938
$a = 1 \times 10^{-7}$	$b = 1$	0.02854
$a = 1 \times 10^{-9}$	$b = 1$	0.04427

Table 2: Values of slope for each year when comparing CLM-Hydro to CLM-Control as shown in Figure 20. Slopes are determined by linear regression between $\log_{10}$(Contributing Area) values of 0.7 to 2.4. (CV* = coefficient of variation)

Variable	2005	2006	2007	2008	2009	2010	Mean	CV*
Evapotranspiration Flux	0.1025	0.0882	0.0757	0.0897	0.1355	0.1058	0.0996	0.2076
Ground Surface Temperature	-0.0022	-0.0020	-0.0017	-0.0021	-0.0028	-0.0020	-0.0021	0.1654
Near Surface Soil Moisture	0.2197	0.1919	0.1818	0.1996	0.2577	0.2064	0.2095	0.1283
Root Zone Soil Moisture	0.2570	0.2335	0.2193	0.2476	0.3176	0.2485	0.2539	0.1336
Soil Temperature	-0.0039	-0.0037	-0.0033	-0.0038	-0.0047	-0.0036	-0.0038	0.1204
Vegetation Water Content	0.0891	0.0660	0.0660	0.0852	0.0917	0.1055	0.0839	0.1844
Water Table Depth	-0.2561	-0.2406	-0.1715	-0.1596	-0.1843	-0.2433	-0.2092	0.2012

FIGURES

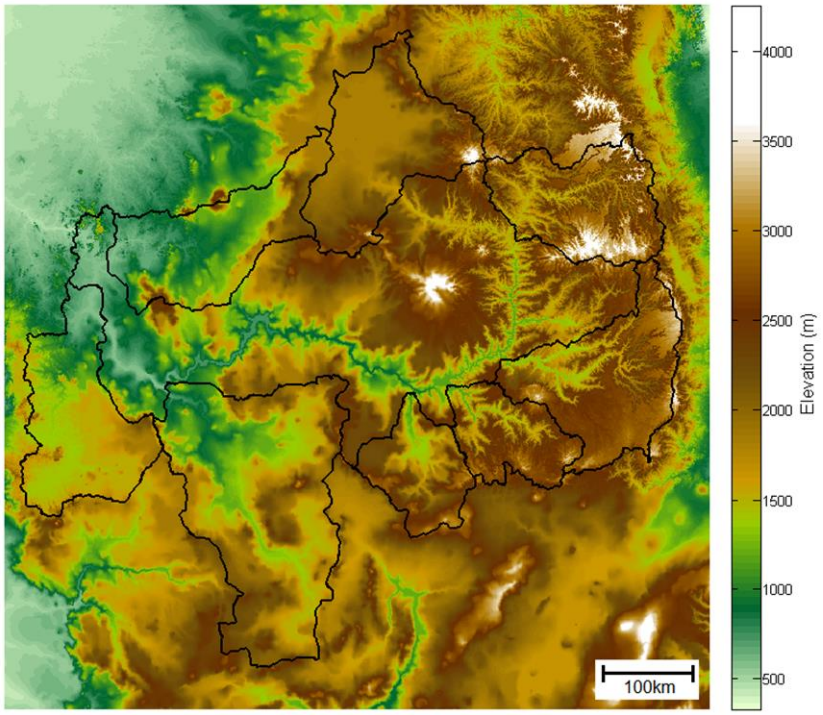

Figure 1: Blue Nile study region exhibiting complex terrain and large river network

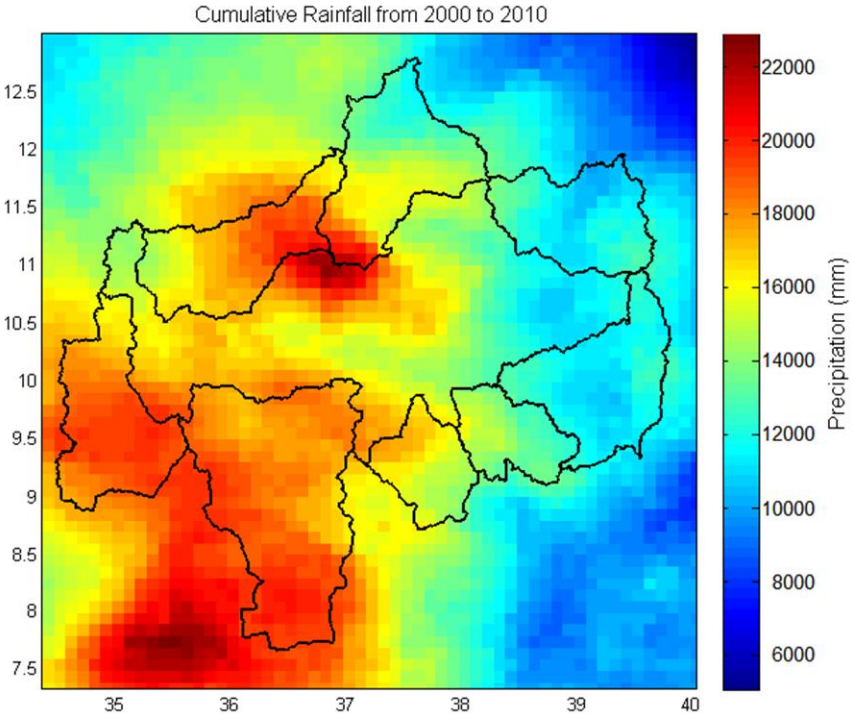

Figure 2: Cumulative Precipitation Contour from 2000 to 2010 using NMA Precipitation

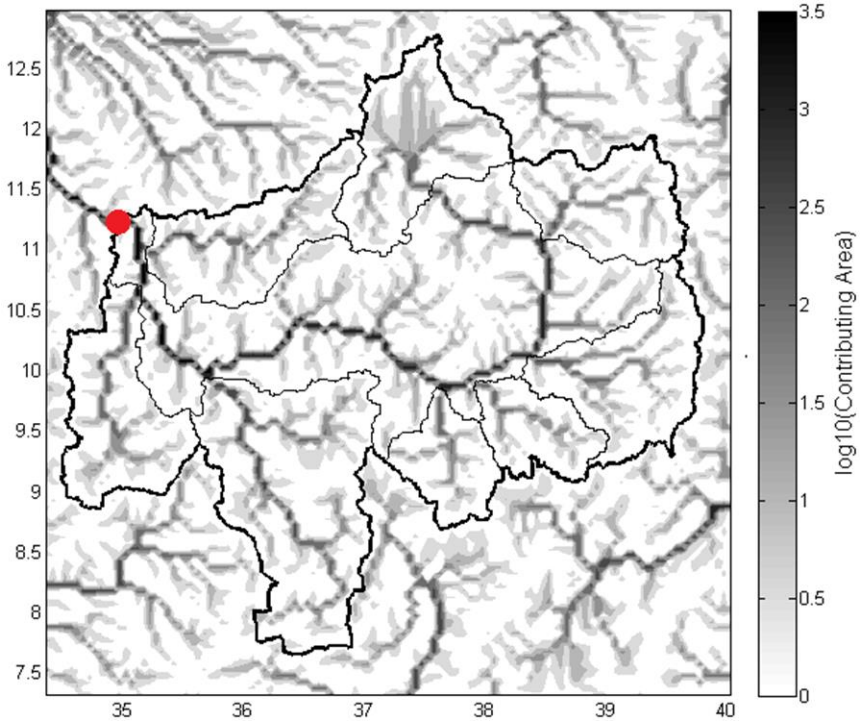

Figure 3: Logarithmic scale figure depicting the contributing area of the region, with Blue Nile Basin outlet identified with red circle

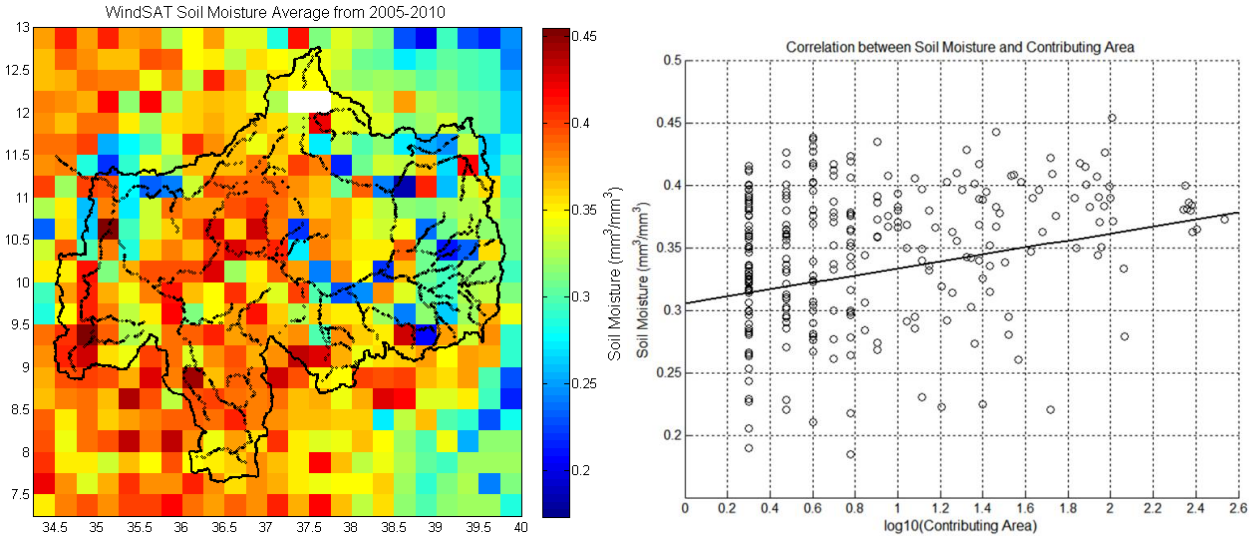

Figure 4: [Left] Contour of NASA WindSAT soil moisture values averaged from 2005 to 2010, river network shown in black. [Right] Plotting soil moisture values against contributing area to highlight any correlation between the two.

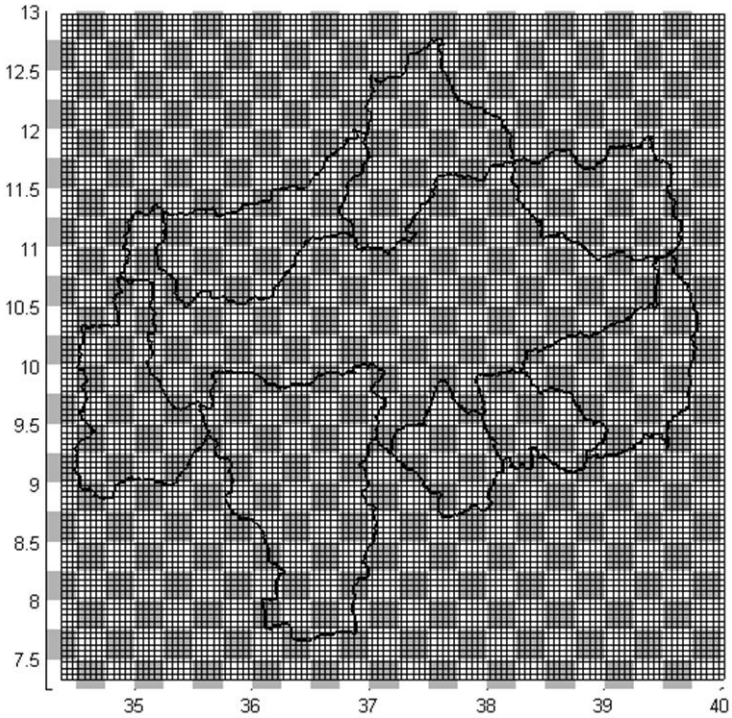

Figure 5: 0.05° resolution mesh (black grid) and 0.025° resolution mesh (gray and white squares) overlaid on Blue Nile Basin boundaries

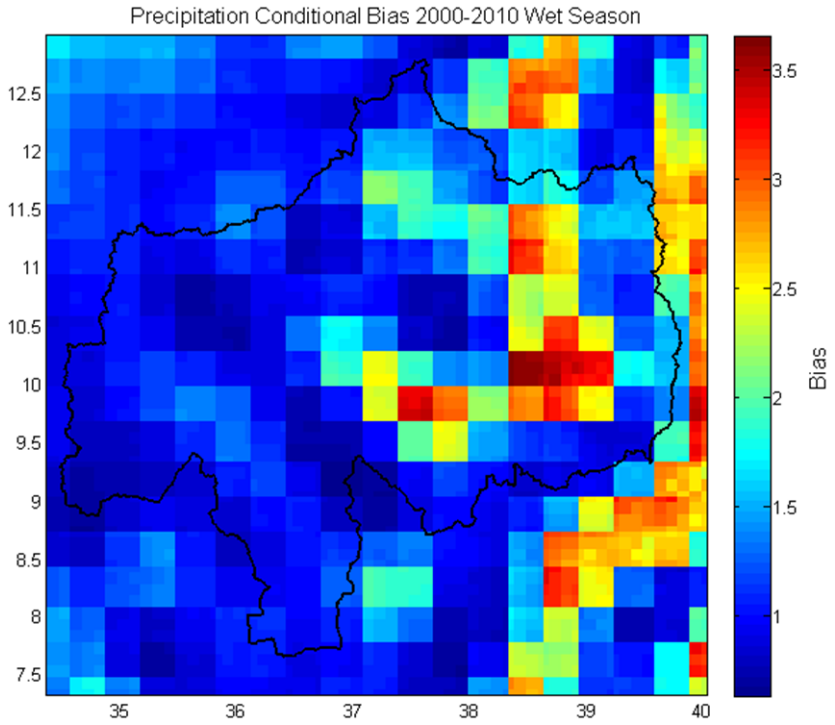

Figure 6: Conditional bias of NCAR CFSR Precipitation Dataset when compared to NMA Precipitation Dataset during months of June, July, August, and September for years of 2000 to 2010

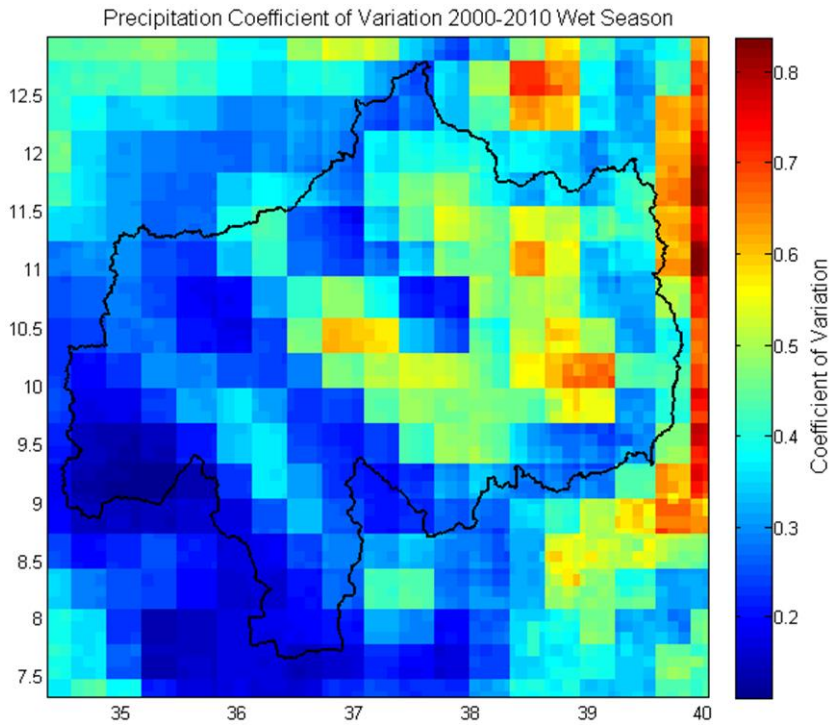

Figure 7: Coefficient of variation for NCAR CFSR Precipitation Dataset when compared to NMA Precipitation Dataset during months of June, July, August, and September for years of 2000 to 2010

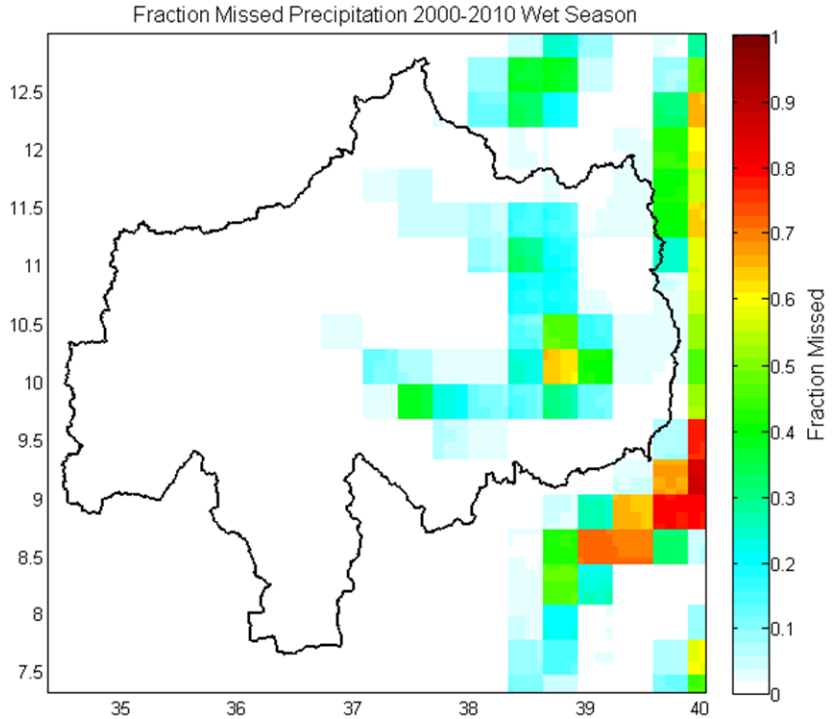

Figure 8: Fraction missed precipitation for NCAR CFSR Precipitation Dataset when compared to NMA Precipitation Dataset during months of June, July, August, and September for years of 2000 to 2010

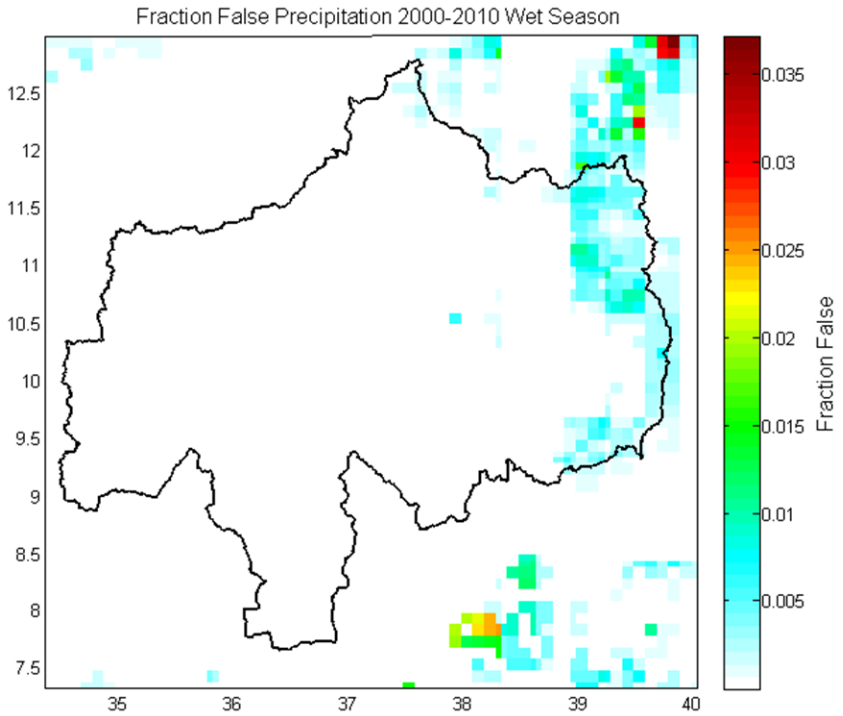

Figure 9: Fraction false precipitation for NCAR CFSR Precipitation Dataset when compared to NMA Precipitation Dataset during months of June, July, August, and September for years of 2000 to 2010

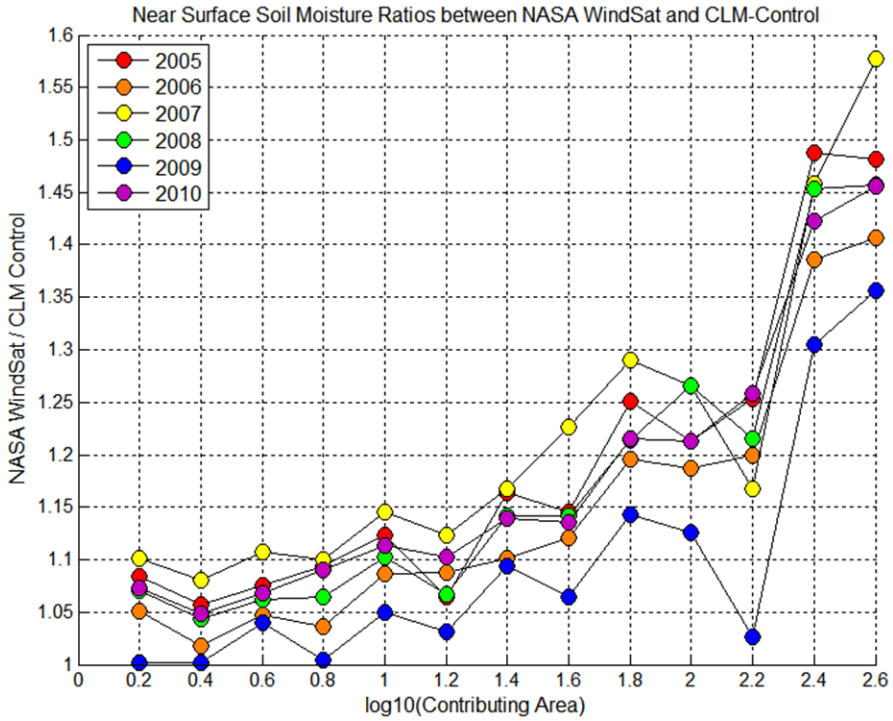

Figure 10: Near-surface soil moisture ratios of NASA WindSat data compared to CLM-Control data to examine model discrepancy at large contributing areas

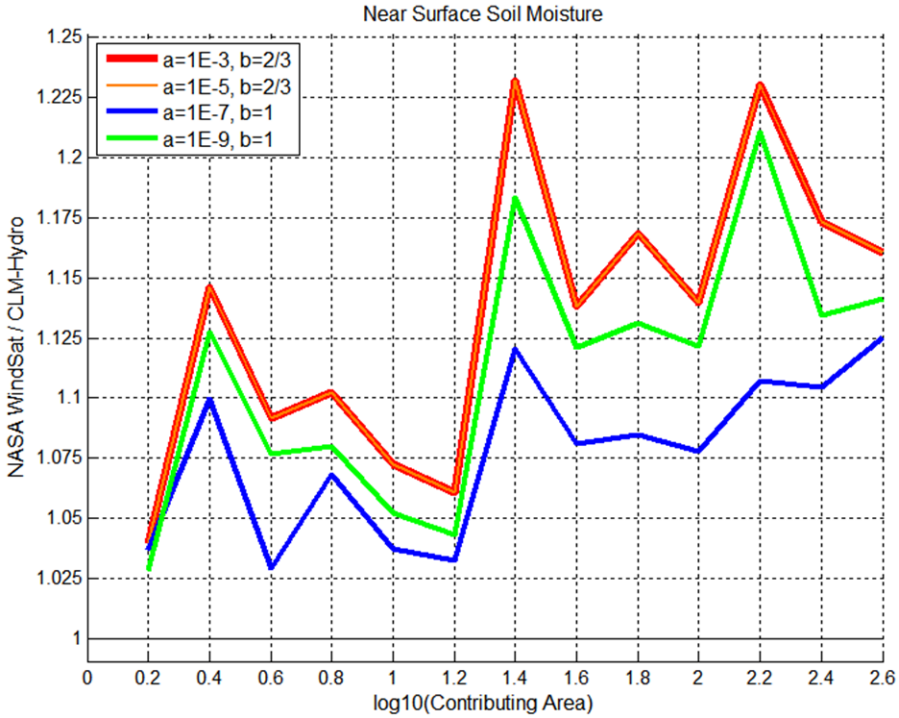

Figure 11: Different parameterizations of CLM-Hydro compared to WindSAT soil moisture dataset, average of years 2005 to 2010

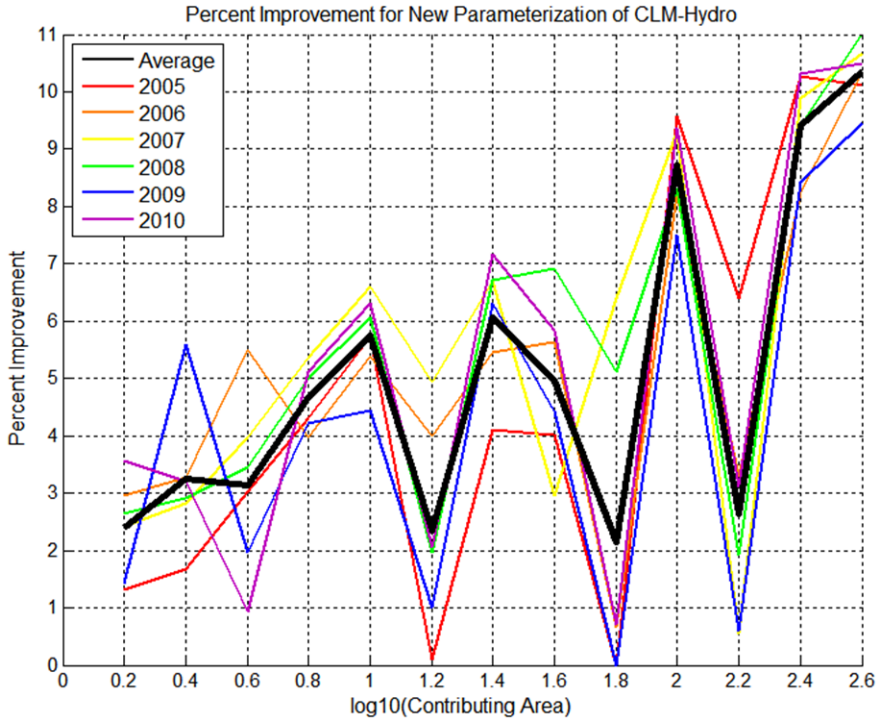

Figure 12: Percent improvement of CLM-Hydro near-surface soil moisture with new parameterization when compared to CLM-Control results from Figure 10

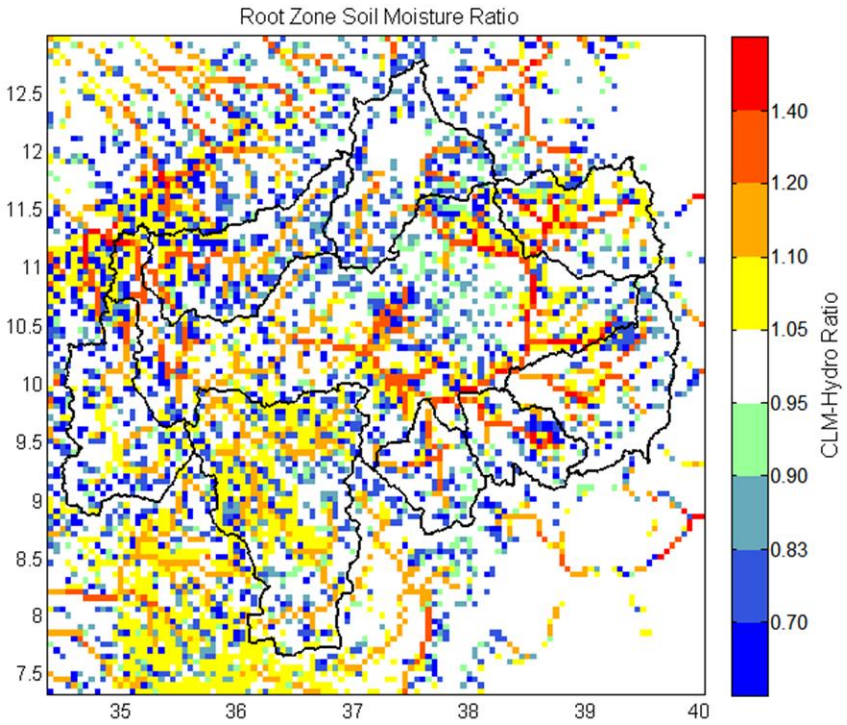

Figure 13: CLM-Hydro and CLM-Control ratio of root zone soil moisture when averaged from 2005 to 2010

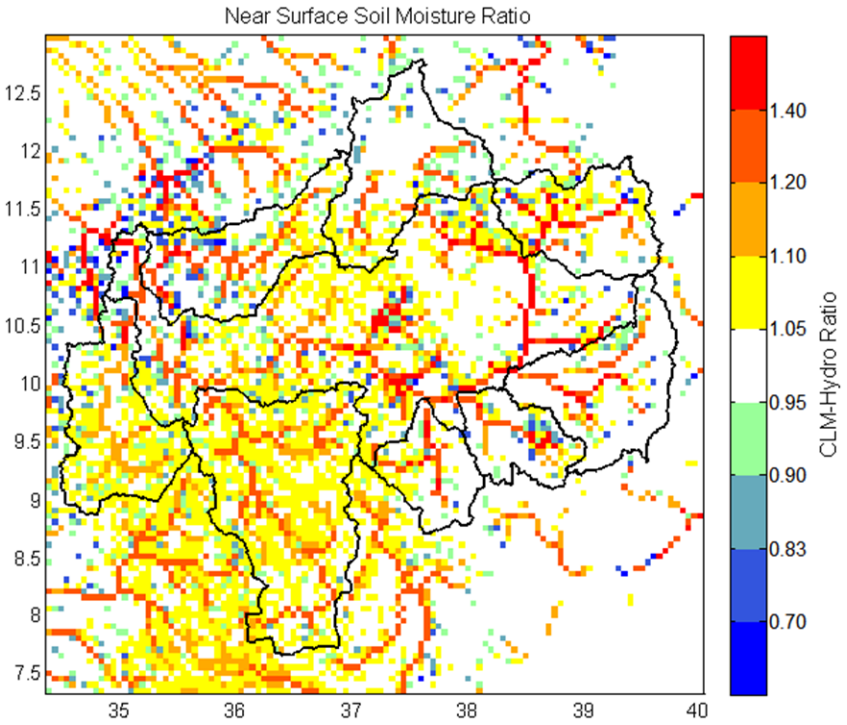

Figure 14: CLM-Hydro and CLM-Control ratio of near-surface soil moisture when averaged from 2005 to 2010

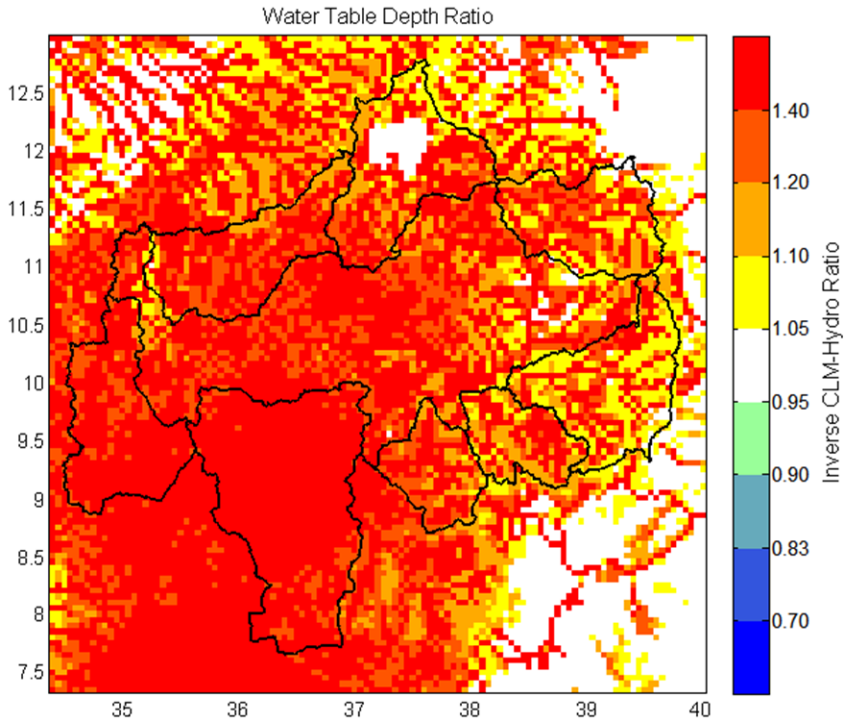

Figure 15: CLM-Hydro and CLM-Control ratio of water table depth when averaged from 2005 to 2010

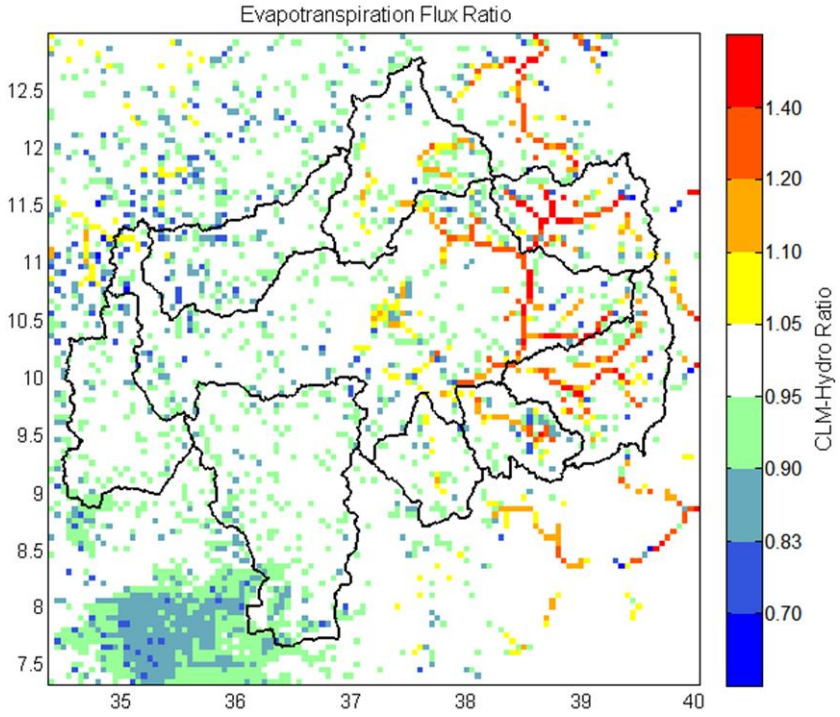

Figure 16: CLM-Hydro and CLM-Control ratio of evapotranspiration flux when averaged from 2005 to 2010

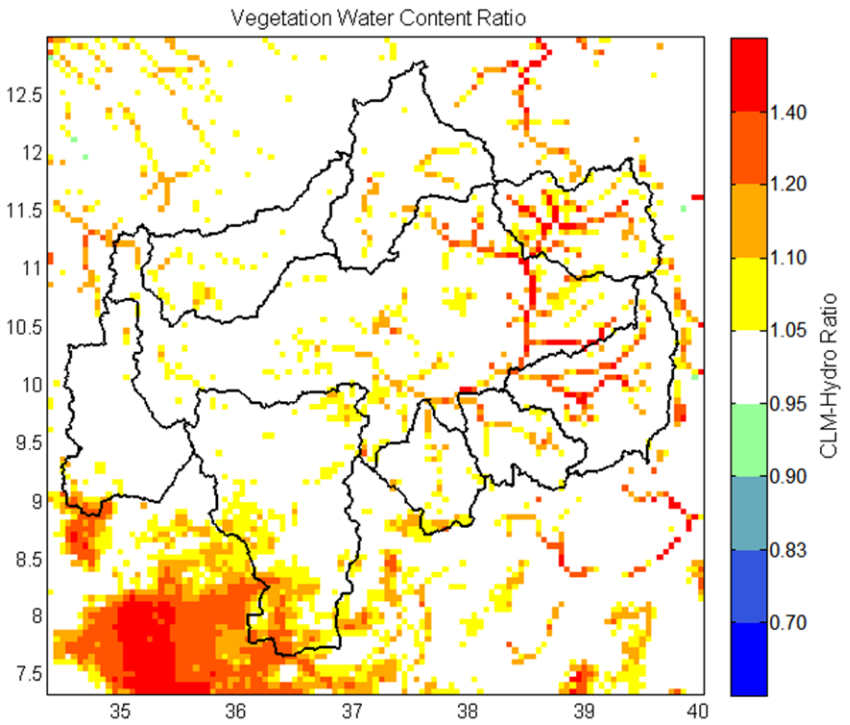

Figure 17: CLM-Hydro and CLM-Control ratio of vegetation water content when averaged from 2005 to 2010

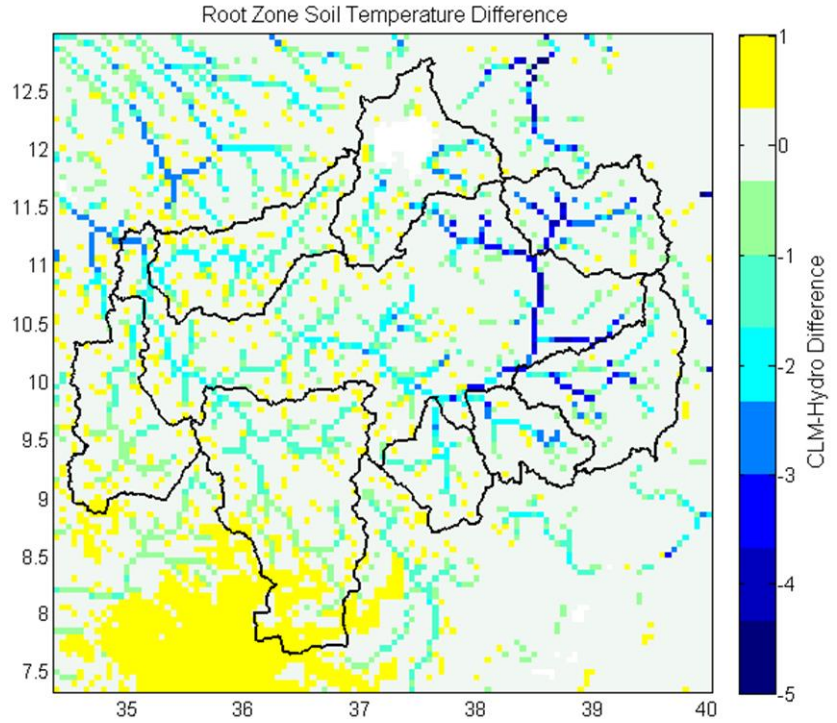

Figure 18: CLM-Hydro and CLM-Control ratio of root zone soil temperature when averaged from 2005 to 2010

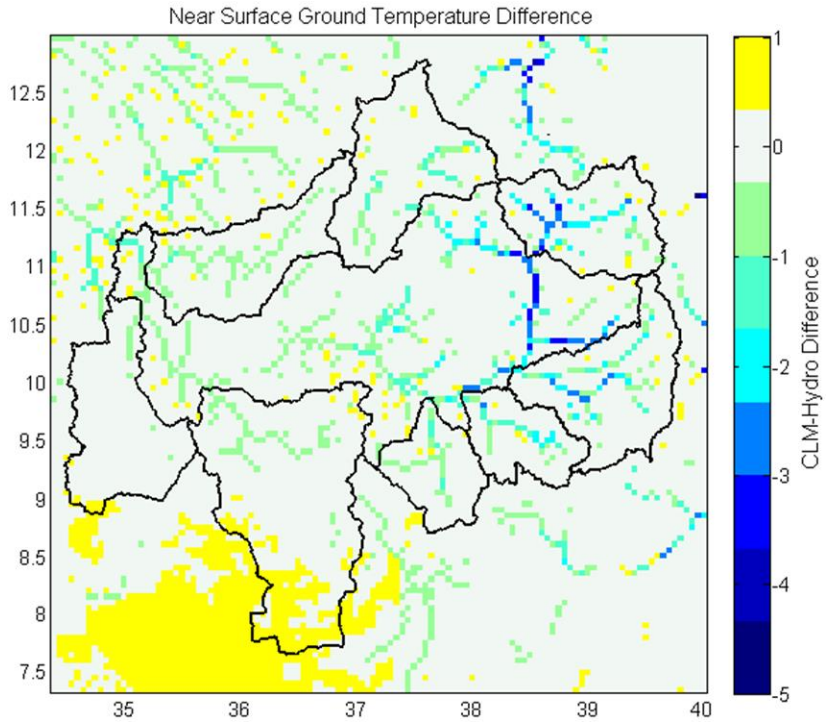

Figure 19: CLM-Hydro and CLM-Control ratio of near-surface ground temperature when averaged from 2005 to 2010

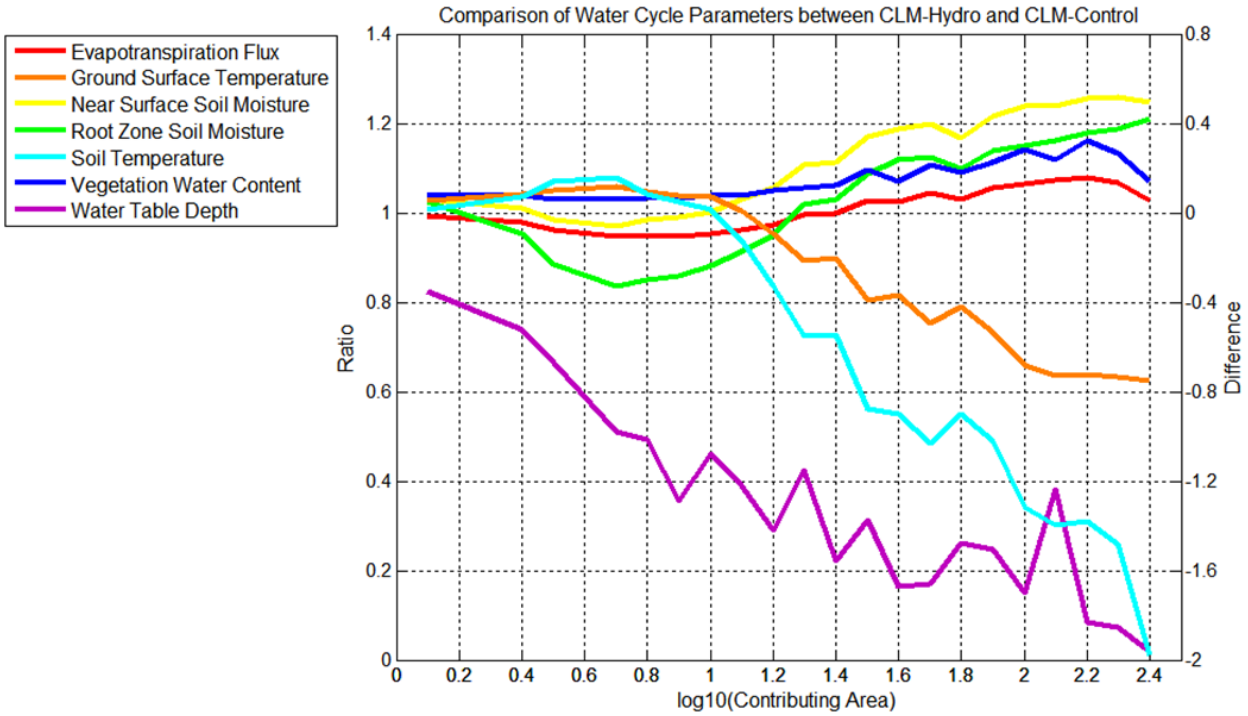

Figure 20: Comparison of water cycle parameters when using the parameterization of surface-groundwater interaction in CLM-Hydro, compared to CLM-Control

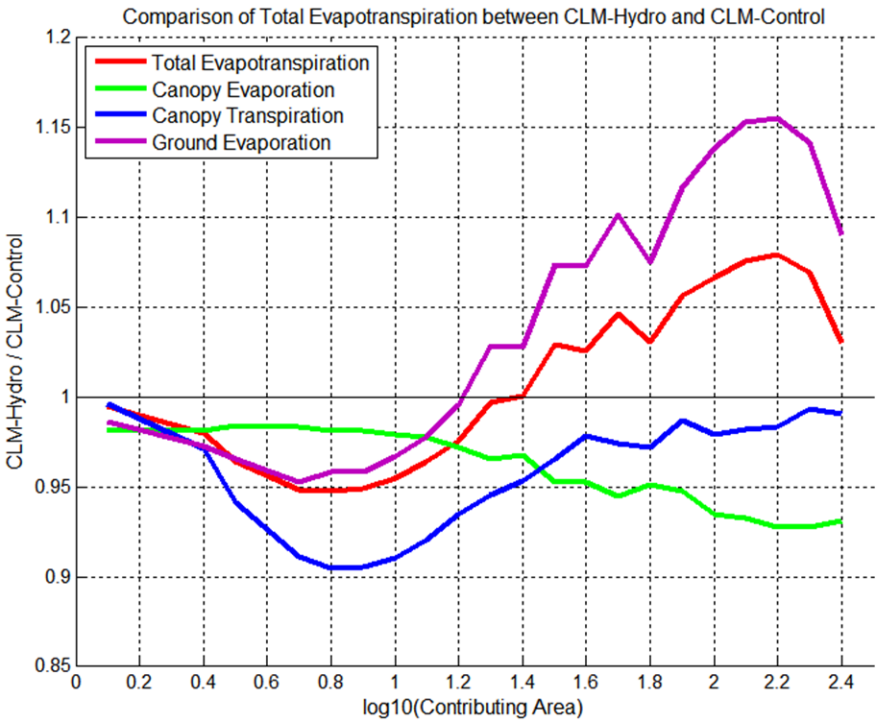

Figure 21: Separation of evapotranspiration component, shown in Figure 20, into canopy evaporation, canopy transportation, and ground evaporation. Averaged ratio between CLM-Hydro and CLM-Control from 2005 to 2010